M. Ram Murty · V. Kumar Murty

The Mathematical Legacy of Srinivasa Ramanujan

Springer

The Mathematical Legacy of Srinivasa Ramanujan

M. Ram Murty • V. Kumar Murty

The Mathematical Legacy of Srinivasa Ramanujan

Springer

M. Ram Murty
Department of Mathematics and Statistics
Queen's University
Kingston, Ontario
Canada

V. Kumar Murty
Department of Mathematics
University of Toronto
Toronto, Ontario
Canada

ISBN 978-81-322-0769-6 ISBN 978-81-322-0770-2 (eBook)
DOI 10.1007/978-81-322-0770-2
Springer New Delhi Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012949894

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

On a height he stood that looked towards greater heights.
Our early approaches to the Infinite
Are sunrise splendours on a marvellous verge
While lingers yet unseen the glorious sun.
What now we see is a shadow of what must come.

Sri Aurobindo, Savitri *1.4*

How I wish I could show you the world through my eyes.

Vivekananda

Preface

22 December 2012 marks the 125th birth anniversary of the Indian mathematician Srinivasa Ramanujan. Being largely self-taught, he emerged from extreme poverty to become one of 20th century's most influential mathematicians. His story is a phenomenal "rags to mathematical riches" story. In his short life, he had a wealth of ideas that have transformed and reshaped 20th century mathematics. These ideas continue to shape mathematics of the 21st century.

This book is meant to be a panoramic view of his essential mathematical contributions. It is not an encyclopedic account of Ramanujan's work. Rather, it is an informal account of some of the major developments that emanated from his work in the 20th and 21st centuries. The twelve essays focus on a subset of his significant papers and show how these papers shaped the course of modern mathematics.

These essays are based on lectures given by the authors over the years at the Chennai Mathematical Institute, Harish-Chandra Research Institute, IISER (Kolkata), IISER (Bhopal), IIT (Powai), IIT (Chennai), Institute for Mathematical Sciences (Chennai), and the Tata Institute for Fundamental Research (Mumbai) as well as Queen's University, the Fields Institute, and the University of Toronto. The lectures were given so that the material is accessible to undergraduates and graduate students. We have striven to not be too technical. At the same time, we tried to convey some depth of the mathematical theories emerging from the work of Ramanujan. Surely, it is impossible to be comprehensive in such a mammoth task. Still, we hope that the reader will see how the vast landscape of Ramanujan's garden has blossomed over the past century.

Toronto, Canada

M. Ram Murty
V. Kumar Murty

Contents

Chapter 1
The Legacy of Srinivasa Ramanujan

Mathematics enjoys the freedom of art and the precision of science. There is freedom of combination of ideas and concepts, but there is also the precision of logic and the ring of truth. It is like a master symphony. The Soviet mathematician, I.R. Shafarevich [186] once remarked that "a superficial glance at mathematics may give an impression that it is a result of separate individual efforts of many scientists scattered about in continents and in ages. However, the inner logic of its development reminds one much more of the work of a single intellect, developing its thought systematically and consistently using the variety of human individualities only as a means. It resembles an orchestra performing a symphony composed by someone. A theme passes from one instrument to another, it is taken up by another and performed with irreproachable precision."

This is no doubt true and yet, the music reaches a crescendo in the hands of certain luminaries. One such luminary was Srinivasa Ramanujan. What is fascinating about Ramanujan is that he was largely self-taught and emerged from extreme poverty to become one of the 20th century's influential mathematicians. His story is a "rags to mathematical riches" story. In the cosmic symphony of mathematics, he played a major role.

The music of Ramanujan emanates both from his life and his work. Born on 22 December 1887 in humble and poor surroundings in the town of Erode situated in present day Tamil Nadu, India, Ramanujan cultivated his love for mathematics singlehandedly and in total isolation. As a child, he was quiet and often to himself. Those that knew him were impressed by his shining large eyes which were his most prominent features. He had a prodigious memory, and at school, he would entertain his friends by reciting the various declensions of Sanskrit roots and by repeating the value of the constant π to any number of decimal places.

At the age of 12, he borrowed a book on trigonometry from an older student and completely mastered its contents. This book was Loney's *Plane Trigonometry* published by Cambridge in 1894 and contains a great deal of information on summation of series, logarithms of complex numbers, calculation of π and Gregory's series. This certainly goes far beyond any modern curriculum of trigonometry taught in our high schools today. But the book that influenced him the most was Carr's *A synopsis*

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_1, © Springer India 2013

of elementary results in pure and applied mathematics. This book is a compilation of some 6165 theorems, systematically arranged but with practically no proofs. It is not a remarkable book and was used by students of Carr for their preparation for the Mathematical Tripos, the entrance examination of Cambridge University. But Ramanujan has made the book famous in that he set about demonstrating to himself each of the assertions enunciated therein. To do this, he used a slate, jotting down the formula to be proved, erasing it with his elbow, jotting down some more formulas that led to the proof, then erasing them again with his elbow and jotting down some more formulas. People used to speak about his bruised elbow, and we know how he got it. Thus he worked his way through the book. This experience influenced him profoundly, and his contact with this book marks the beginning of his exploration of the world of mathematics. Carr's synopsis was therefore a great blessing. But unfortunately, Ramanujan took this synopsis as his model for writing, and his famous notebooks consisting of over 4000 formulas are written down in this style without proofs. The intermediate results, the links of the chain, have been erased by the elbow of Ramanujan, and his legacy is simply a set of discoveries, a melody of formulas.

When we look through these formulas discovered by Ramanujan, it is like great music echoing through our consciousness and the music lingers. Each is pregnant with meaning and heralding further exploration. Perhaps it is not so unfortunate that Ramanujan had taken Carr as his model for writing. We all now have work to do. "When the kings are building, the carters have work to do."

Another thing that we learn from the early mathematical development of Ramanujan, is the importance of problem solving in the primary grades. The mathematician and educator, George Pólya, was right when he stated that mathematics cultivates logical and orderly thinking, and a precision for the expression of ideas. So Ramanujan mastered a large tract of college level mathematics simply through problem solving and working through Carr's synopsis.

A year later, in 1903, he secured a seat in the Government College in Kumbakonam. However, his passionate absorption in mathematics led him to neglect his other subjects, and the inevitable happened. He failed the exams at the end of his first year. Four years later, he entered another college in Madras (now called Chennai) but met with the same fate at the end of his first year.

In 1909, at the age of 22, he married Kumari Janaki, and with his new responsibility, it was necessary for him to secure a job. This he succeeded in doing in 1912, when he became a clerk in the Madras Port Trust Office. There his duties were light, and he found time to devote to his mathematical research. Moreover, as luck would have it, the manager of the office, S.N. Aiyar, was a mathematician who took kindly to him and his discoveries. With Aiyar's encouragement, Ramanujan communicated some of his results to several British mathematicians. (For a short biography of S.N. Aiyar and the role he played in Ramanujan's life, we refer the reader to a recent article by Berndt [23].) His first three attempts produced little or no response. But in 1913, he wrote to G.H. Hardy at Trinity College, Cambridge. This was a turning point since Hardy was a renowned expert in analysis and number theory.

We should say here that number theory should not be confused with numerology. There is no mysticism attached to number theory. The only mystifying element is that there are beautiful formulas and there is a logical, mathematical order in the apparent chaotic universe. Number theory is the study of hidden mathematical patterns among numbers. It is called the queen of mathematics because the problems of number theory have given birth to the diverse disciplines of mathematics. Problems are utilized as points of focus of concentration. In themselves, the problems are unimportant. But in the finding of their solution, new concepts arise, and new links and patterns are found with other concepts and disciplines of mathematics. It is the final mosaic that is the end in view and not the esoteric problem, which is used only as a means of motivation.

So when Hardy received the letter, he found himself a little confounded and could not at first decide whether it was written by a crank or a genius. To the letter were attached about 120 theorems of which a representative sample is given by the following 15:

$$1-\frac{3!}{(1!2!)^3}x^2+\frac{6!}{(2!4!)^3}x^4-\cdots = \left(1+\frac{x}{1!^3}+\frac{x^2}{2!^3}+\cdots\right)\left(1-\frac{x}{1!^3}+\frac{x^2}{2!^3}-\cdots\right). \tag{1}$$

$$1-5\left(\frac{1}{2}\right)^3+9\left(\frac{1\cdot 3}{2\cdot 4}\right)^3-13\left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^3+\cdots=\frac{2}{\pi}. \tag{2}$$

$$1+9\left(\frac{1}{4}\right)^4+17\left(\frac{1\cdot 5}{4\cdot 8}\right)^4+25\left(\frac{1\cdot 5\cdot 9}{4\cdot 8\cdot 12}\right)^4+\cdots=\frac{2^{3/2}}{\sqrt{\pi}\Gamma(3/4)^2}. \tag{3}$$

$$1-5\left(\frac{1}{2}\right)^5+9\left(\frac{1\cdot 3}{2\cdot 4}\right)^5-13\left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^5+\cdots=\frac{2}{\Gamma(3/4)^4}. \tag{4}$$

$$\int_0^\infty \frac{1+(\frac{x}{b+1})^2}{1+(\frac{x}{a})^2}\cdot\frac{1+(\frac{x}{b+2})^2}{1+(\frac{x}{a+1})^2}\cdots dx = \frac{\sqrt{\pi}}{2}\frac{\Gamma(a+\frac{1}{2})\Gamma(b+1)\Gamma(b-a+\frac{1}{2})}{\Gamma(a)\Gamma(b+\frac{1}{2})\Gamma(b-a+1)}. \tag{5}$$

$$\int_0^\infty \frac{dx}{(1+x^2)(1+r^2x^2)(1+r^4x^2)\ldots}=\frac{\pi}{2(1+r+r^3+r^6+r^{10}+\cdots)}. \tag{6}$$

If $\alpha\beta=\pi^2$, then

$$\alpha^{-1/4}\left(1+4\alpha\int_0^\infty \frac{xe^{-\alpha x^2}}{e^{2\pi x}-1}dx\right)=\beta^{-1/4}\left(1+4\beta\int_0^\infty \frac{xe^{-\beta x^2}}{e^{2\pi x}-1}dx\right). \tag{7}$$

$$\int_0^a e^{-x^2}dx = \frac{1}{2}\sqrt{\pi} - \frac{e^{-a^2}}{2a+}\frac{1}{a+}\frac{2}{2a+}\frac{3}{a+}\frac{4}{2a+\cdots}. \tag{8}$$

$$4\int_0^\infty \frac{xe^{-x\sqrt{5}}}{\cosh x}dx = \frac{1}{1+}\frac{1^2}{1+}\frac{1^2}{1+}\frac{2^2}{1+}\frac{2^2}{1+}\frac{3^2}{1+}\frac{3^2}{1+\cdots}. \tag{9}$$

$$\text{If}\quad u = \frac{x}{1+}\frac{x^5}{1+}\frac{x^{10}}{1+}\frac{x^{15}}{1+\cdots},\quad v = \frac{x^{1/5}}{1+}\frac{x}{1+}\frac{x^2}{1+}\frac{x^3}{1+\cdots},$$
$$\text{then}\quad v^5 = u\frac{1-2u+4u^2-3u^3+u^4}{1+3u+4u^2+2u^3+u^4}. \tag{10}$$

$$\frac{1}{1+}\frac{e^{-2\pi}}{1+}\frac{e^{-4\pi}}{1+\cdots} = \left\{\sqrt{\left(\frac{5+\sqrt{5}}{2}\right)} - \frac{\sqrt{5}+1}{2}\right\}e^{2\pi/5}. \tag{11}$$

$$\frac{1}{1+}\frac{e^{-2\pi\sqrt{5}}}{1+}\frac{e^{-4\pi\sqrt{5}}}{1+\cdots} = \left[\frac{\sqrt{5}}{1+(5^{3/4}(\frac{\sqrt{5}-1}{2})^{5/2}-1)^{1/5}} - \frac{\sqrt{5}+1}{2}\right]e^{2\pi/\sqrt{5}} \tag{12}$$

$$\text{If}\quad F(k) = 1+\left(\frac{1}{2}\right)^2 k + \left(\frac{1\cdot 3}{2\cdot 4}\right)^2 k^2+\cdots \quad\text{and}$$
$$F(1-k) = \sqrt{210}F(k),\quad \text{then}$$
$$k = (\sqrt{2}-1)^4(2-\sqrt{3})^2(\sqrt{7}-\sqrt{6})^4(8-3\sqrt{7})^2(\sqrt{10}-3)^4$$
$$\times(4-\sqrt{15})^4(\sqrt{15}-\sqrt{14})^2(6-\sqrt{35})^2. \tag{13}$$

The coefficient of x^n in $(1-2x+2x^4-2x^9+\cdots)^{-1}$ is the integer nearest to

$$\frac{1}{4n}\left(\cosh\pi\sqrt{n} - \frac{\sinh\pi\sqrt{n}}{\pi\sqrt{n}}\right). \tag{14}$$

The number of numbers between A and x which are either squares or sums of two squares is

$$K\int_A^x \frac{dt}{\sqrt{\log t}} + \theta(x), \tag{15}$$

where $K = 0.764\ldots$, and $\theta(x)$ is very small compared with the previous integral.

These fifteen entries from Ramanujan's letter to Hardy give a representative sample of the formulas contained there. The first four belong to the theory of infinite series. The next three are new definite integrals. Formulas (8) to (12) are in the theory of continued fractions. Formula (13) belongs to the theory of complex multiplication and singular moduli. Formula (14) is the first suggestion of Ramanujan's knowledge

of the circle method (about which we say more in a later chapter). Finally, (15) belongs to analytic number theory.

Hardy took over two hours to analyse the letter to ascertain whether the author was a crank or a genius. Hardy's reaction is expressed in his own words: "I should like you to begin by trying to reconstruct the immediate reactions of an ordinary professional mathematician who receives a letter like this from an unknown Indian clerk."

"The first question was whether I could recognise anything. I had proved things rather like (7) myself and seemed vaguely familiar with (8). Actually (8) is classical; it is a formula of Laplace first proved properly by Jacobi and (9) occurs in a paper published by Rogers in 1907."

So the conclusion was that Ramanujan had rediscovered all of these theorems amidst the impoverished mathematical background of his rustic surroundings.

Hardy continues, "I thought, that as an expert in definite integrals, I could probably prove (5) and (6) and did so, though with a good deal more trouble than I had expected. ... The series formulas (1)–(4) I found much more intriguing and it soon became obvious that Ramanujan must possess much more general theorems and was keeping a great deal up his sleeve. ... The formulas (10)–(13) are on a different level and obviously both difficult and deep. An expert in elliptic functions can see at once that (13) is derived somehow from the theory of complex multiplication, but (10)–(12) defeated me completely; I had never seen anything in the least like them before. ... The last two formulas stand apart. ... The function in (14) is a genuine approximation to the coefficient, though not at all close as Ramanujan imagined and Ramanujan's false statement was one of the most fruitful he ever made, since it ended by leading us to all our joint work on partitions."

Indeed, (14) could only be derived by the circle method, a powerful technique developed later by Hardy and Ramanujan in their work on the partition function. The entry in the letter shows that Ramanujan had already thought about the circle method in India, before he had met Hardy.

It seems that Hardy invited his colleague Littlewood and showed him the letter. They sat with it for three hours, from 9pm to midnight and finally concluded that this indeed was the work of a genius. Hardy [67] wrote later, "A single look at them is enough to show that they could only be written down by a mathematician of the highest class. They must be true, because if they were not true, no one would have had the imagination to invent them."

Soon thereafter, Hardy invited Ramanujan to come to Cambridge, which he felt could provide a better environment in which his mathematical genius could flourish. So Ramanujan sailed for England in March 1914.

Against the background of the first world war, from 1914 to 1917, Hardy and Ramanujan spent time in wonderful mathematical collaboration. Hardy remarks that every day Ramanujan would show him about half a dozen new theorems. These three years saw prodigious mathematical activity by both Hardy and Ramanujan. In that period, Ramanujan wrote over 30 papers, in which were laid the foundations of three fundamental methods in number theory.

The first of them concerns the circle method which, as noted earlier, already has its genesis in Eq. (14) of his first letter to Hardy. This method concerns an

ingenious idea for computing the integral of a function by studying its behaviour at rational points and sufficiently small neighborhoods. The method can be developed further and enables one to attack classical unsolved problems such as Goldbach's conjecture, Waring's problem and explicit formulas for the Fourier coefficients of modular forms and modular functions. These problems have defied solution for a long time. Due to Ramanujan's early demise, this work was carried on by Hardy and Littlewood, and today it is called the circle method, or the Hardy–Littlewood method. In their paper on the partition function, Hardy and Ramanujan laid the groundwork of the method. But it was clear that this was a viable technique for attacking many age-old problems. Subsequently, this technique was developed and improved by Hardy and Littlewood and I.M. Vinogradov.

The second major contribution was the normal order method which proved that almost all natural numbers have $\log\log n$ prime factors. These investigations were later developed into the beautiful probabilistic theory of numbers, starting with the work of Turán and culminating in the celebrated Erdös–Kac theorem. Afterwards, it was taken up by Kubilius, who infused finer improvements into the theory. We feel that there is a further direction for these investigations into the domain of Fourier coefficients of modular forms, and we discuss this theme in this monograph. This brings us to the third great contribution of Ramanujan.

In 1916, he wrote a classic paper entitled "On certain arithmetical functions" in which he investigated the Fourier coefficients of various modular forms. There, he noticed patterns of congruences and made three significant conjectures concerning the behaviour of these Fourier coefficients. The most famous of these concerns the Ramanujan τ-function. He conjectured that this function satisfies a multiplicative law and that its growth is controlled by a simple polynomial function. At the time, the conjecture did not have much meaning other than as an esoteric problem in analytic number theory. Indeed, regarding the τ-function, Hardy [67] wrote, "We may seem to be straying into one of the backwaters of mathematics, but the genesis of $\tau(n)$ as a coefficient in so fundamental a function compels us to treat it with respect."

Putting respect aside, the fact is that the function occupies a central place in the pantheon of coefficients of modular forms. So the legacy left by Ramanujan's conjecture is vast and deep. For it slowly transpired that these conjectures had intimate connection with profound aspects of number theory and algebraic geometry.

The first significant step towards the conjectures of Ramanujan was taken by Mordell, who proved the multiplicative law, the first part of Ramanujan's conjectures. But Mordell only treated the case of τ and did not realize that it was prototypical of a spectrum of functions, each in its own right of central importance. This realization came twenty years later, in the work of the German mathematician Erich Hecke (who incidently was also born in the same year as Ramanujan), and the meaning of the multiplicative properties conjectured by Ramanujan was unravelled. This work of Hecke is considered a masterpiece of mathematics.

But the connections to algebraic geometry were deeper still. Indeed, after the pioneering work of Artin and Hasse, Weil formulated in 1949 general conjectures

about solutions of systems of equations over finite fields. In the 1950s, it was suspected that the τ-function of Ramanujan enumerates the number of these solutions for a certain system of equations (called a variety). Several Japanese mathematicians constructed the possible candidate, but there were technical problems related to compactification. These problems were resolved in 1974 by a Belgian mathematician Pierre Deligne, who completely settled the conjectures of Ramanujan. In 1976, Deligne was awarded the Fields medal (which is the mathematical analogue of the Nobel prize) for this achievement. This is quite a legacy!

The connection between Ramanujan's conjecture and Weil's conjecture could not be so easily foreseen. Both conjectures reflect elements of the master symphony that Shafarevich spoke about. Indeed, in his retrospective essay on number theory, André Weil wrote [201], "In 1947, in Chicago, I felt bored and depressed, and not knowing what to do, I started reading Gauss's two memoirs on biquadratic residues, which I had never read before. The Gaussian integers occur in the second paper. The first one deals essentially with the number of solutions of equations $ax^4 - by^4 = 1$ in the prime field modulo p, and with the connection between these and certain Gaussian sums. ... then I noticed that similar principles can be applied to all equations of the form $ax^m + by^n + cz^r + \cdots = 0$, and that this implies the truth of the so-called "Riemann hypothesis" for all curves $ax^n + by^n + cz^n = 0$ over finite fields, and also a "generalized Riemann hypothesis" for varieties in projective space with a "diagonal" equation $\sum a_i x_i^n \equiv 0$. This led me in turn to conjectures about varieties over finite fields."

It was only a matter of time before several notable mathematicians realized that Ramanujan's conjecture was really a "Riemann hypothesis" for a certain zeta function of a variety over a finite field and that it would follow from Weil's conjecture. This was the achievement of Pierre Deligne in 1974.

No one could have foreseen such a cosmic connection. Yet, Weil is quite harsh on Hardy and in the same essay [201] wrote "Hardy's remarkable comment is: "We seem to have drifted into one of the backwaters of mathematics." To him it was just another inequality; he found it curious that anyone could get deeply interested in it. In fact, he becomes apologetic and explains that, in spite of the apparent lack of interest of this problem it might still have some features which made it not unworthy of Ramanujan's attention."

The problem with Weil's assessment of Hardy is that it is inaccurate. Hardy's original quotation is that "We may seem to be straying into one of the backwaters of mathematics, but the genesis of $\tau(n)$ as a coefficient in so fundamental a function compels us to treat it with respect." The reader will note that Weil replaced "We may seem" with "We seem" which gives quite a twist to the meaning. Moreover, Hardy does not say we will study it simply because Ramanujan had studied it, but rather that it is the coefficient of a fundamental function, namely $\Delta(z)$ in the theory of modular forms. If Hardy thought that the τ-function was in the backwaters of mathematics, then it would have been unreasonable to give it as a doctoral thesis problem to one of his celebrated students, R.A. Rankin, who created what is now called Rankin's method (also called the Rankin–Selberg method) an analogue of which was instrumental in the resolution of the Weil conjectures. So it is

quite presumptuous to make absolute pronouncements on the significance of various mathematical ideas since we never know how ideas are interconnected. And this is part of the legacy.

But this legacy does not end here, and these investigations form just the tip of an iceberg. The central problem of number theory revolves around what is called the reciprocity law. The function of Ramanujan, and Fourier coefficients in general and their congruence properties reflect some aspects of the non-abelian reciprocity law. The theory of modular forms was further generalized by Jacquet and Langlands, and higher-dimensional versions of Ramanujan's conjectures were formulated as part of the Langlands program. Some of these conjectures go beyond the Ramanujan conjecture. Finer distribution conjectures concerning the τ-function inspired by the work of Sato and Tate in the theory of elliptic curves and first enunciated by Serre have now been proved. These results represent Himalayan peaks in the mathematical landscape of the 21st century.

Returning to our narrative of Ramanujan, we find that at the end of his three years work in England, he left behind a tremendous mathematical legacy. In the summer of 1917, he fell ill with what was suspected to be tuberculosis. He never recovered. Nevertheless, he continued to work unabated. Hardy relates an interesting story during the time that Ramanujan was staying in the hospital in Putney. He went to visit him in a taxicab, and as he entered Ramanujan's room, remarked that he had just ridden in a taxicab with number 1729 which seemed to be to him a rather dull number and hoped that this was not the indication of a bad omen. Ramanujan replied that on the contrary it is a very interesting number. It is the smallest number which can be expressed as the sum of two cubes in exactly two different ways:

$$1729 = 1^3 + 12^3 \quad = 10^3 + 9^3.$$

This is not without significance. Recently, a beautiful theorem in the theory of elliptic curves was proved involving this taxicab number. If for each k, there is a squarefree natural number n that can be expressed in at least k different ways as the sum of two cubes, then an outstanding problem in the theory of elliptic curves has an affirmative solution, namely ranks of elliptic curves over $\mathbb{Q}$ tend to infinity.

Ramanujan had great intuition into what was important and central. His facile mind revealed an artistic symbiosis between intellect and inspiration. We are reminded of a statement made by a great Indian sage, Swami Vivekananda, over a century ago. He said: "Just as the intellect is the instrument of knowledge, so is the heart the instrument of inspiration. In a lower state, the heart is a much weaker instrument than the intellect. ... Properly cultivated, the heart can be changed and will go beyond intellect; it will function through inspiration. Man will have to go beyond intellect in the end. The knowledge of man, his powers of perception, of reasoning and intellect and heart, all are busy churning this milk of the world. Out of long churning comes butter ... Men of [cultivated] heart get the butter and the buttermilk is left for the intellectual."

We do not know how Ramanujan discovered his theorems. On this point, Hardy [67] wrote, "It was his insight into algebraical formulas, transformations of infinite

series and so forth, that was most amazing. On this side most certainly I have never met his equal, and I can compare him only with Euler or Jacobi. He worked far more than the majority of modern mathematicians, by induction from numerical examples; all his congruence properties of partitions, for example, were discovered in this way. But with his memory, his patience, and his power of calculation he combined a power of generalisation, a feeling for form, a capacity for rapid modification of his hypothesis, that were often really startling, and made him, in his own peculiar field, without a rival in his day."

These comments were made by Hardy in 1936 when he delivered his famous Harvard lectures on the work of Ramanujan. He began them with a rather sentimental tone. "I have to help you," he said, "to form some sort of reasoned estimate of the most romantic figure in the recent history of mathematics ... Ramanujan was, in a way, my discovery. I did not invent him—like other great men, he invented himself—but I was the first really competent person who had a chance to see some of his work, and I can still remember with satisfaction that I could recognize at once what a treasure I had found ... And my association with him is the one romantic incident in my life." These are powerful feelings indeed describing one of the great collaborations of mathematics!

We conclude this introduction by reflecting upon what we call the cultural legacy left behind by Ramanujan. We can do this no better than to relate the feelings expressed by the Nobel laureate, Subramanian Chandrasekhar, at the Ramanujan Centennial Conference in Urbana in 1987. He wrote [31]: "It must have been a day in April 1920, when I was not quite ten years old, when my mother told me of an item in the newspaper of the day that a famous Indian mathematician, Ramanujan by name, had died the preceding day; and she told me further that Ramanujan had gone to England some years earlier, had collaborated with some famous English mathematicians and that he had returned only very recently, and was well known internationally for what he had achieved. Though I had no idea at that time of what kind of a mathematician Ramanujan was, or indeed what scientific achievement meant, I can still recall the gladness I felt at the assurance that one brought up under circumstances similar to my own, could have achieved what I could not grasp. I am sure that others were equally gladdened. I hope that it is not hard for you to imagine what the example of Ramanujan could have provided for young men and women of those times, beginning to look at the world with increasingly different perceptions.

"The fact that Ramanujan's early years were spent in a scientifically sterile atmosphere, that his life in India was not without hardships, that under circumstances that appeared to most Indians as nothing short of miraculous, he had gone to Cambridge, supported by eminent mathematicians and had returned to India with every assurance that he would be considered, in time, as one of the most original mathematicians of the century—these facts were enough—more than enough—for aspiring young Indian students to break their bonds of intellectual confinement and perhaps soar the way that Ramanujan did.

"It may be argued, perhaps with some justice, that this was a sentimental attitude: Ramanujan represents so extreme a fluctuation from the norm that his being

born an Indian must be considered to a large extent as accidental. But to the Indians of the time, Ramanujan was not unique in the way we think of him today. He was one of others who had, during that same period, achieved, in their judgement, comparably in science and in other areas of human activity. Gandhi, Nehru, Rabindranath Tagore, J.C. Bose, C.V. Raman, M.N. Saha, S.N. Bose and a host of others, were in the forefront of the then fermenting scene."

In these words of Chandrasekhar, we see the legacy of Ramanujan. For the life of Chandrasekhar was equally full of hardships. Born in the same village surroundings as Ramanujan, he went to study at Cambridge and there as a graduate student discovered the mathematical implications of the theory of relativity in the collapse of certain massive stars. These he predicted degenerate into black holes. The high priests of physics of that time rejected his calculations as meaningless. He had to wait for another thirty years before the theory came into the forefront of modern physics and finally in 1983, he was awarded the Nobel prize in physics as recognition of his work. The life of Subramanian Chandrasekhar itself reveals to some extent the grandeur of the legacy of Ramanujan.

But a scientist belongs to no nation. Many of the mathematicians of distinction that we have met and talked with have all told us that Ramanujan directly or indirectly inspired their mathematical life. This is not surprising. For as we have seen, Ramanujan embodies that marvelous miracle of the human mind to frame concepts and to use formulas and symbols as tools of thought to probe deeper into the mysteries of one's own being. As long as the spirit of science is alive, his legacy will live, and the music will pass from one luminary to another. And all of us who think and work with mathematical ideas are participants in that wonderful symphony.

Chapter 2
The Ramanujan τ-Function

1 Introduction

The 1916 memoir of Ramanujan, innocuously entitled "On certain arithmetic functions", introduced the τ-function. This is an integer-valued function on the natural numbers which, at first, manifested as part of an "error term" in counting the number of ways that a number could be written as a sum of 24 squares. However, Ramanujan realized that it was a function worthy of study in its own right. It would not be an overstatement to say that one of the significant themes of mathematics in the 20th century has emanated from this observation.

The τ-function is defined by the formal identity

$$\sum_{n=1}^{\infty} \tau(n) q^n = q \prod_{n=1}^{\infty} \left(1 - q^n\right)^{24}. \tag{1}$$

It is in fact more than a formal identity. If we think of q as a complex number with $|q| < 1$, then taking the logarithm of the infinite product and expanding, we see that it is

$$\log q - 24 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{q^{nm}}{m}.$$

Interchanging sums, we see that the double sum is

$$-24 \sum_{m=1}^{\infty} \frac{1}{m} \frac{q^m}{1 - q^m},$$

and this converges for $|q| < 1$. For a complex number z with $\Im(z) > 0$, we see that

$$q = e^{2\pi i z}$$

satisfies $|q| < 1$, and so we may define a function $\Delta(z)$ by the right-hand side of (1). Moreover, we see that $\Delta(z) \neq 0$ as it is given as an absolutely convergent product of non-zero terms.

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_2,

The function $\Delta(z)$ was known to previous authors. If we consider the Eisenstein series

$$E_{2k}(z) = \sum_{(m,n)\neq(0,0)} \frac{1}{(mz+n)^{2k}}$$

where the sum ranges over pairs of integers (m, n) which are not both zero, then there is the identity

$$\Delta(z) = \frac{1}{1728}\big(E_4(z)^3 - E_6(z)^2\big).$$

Dedekind had studied $\Delta(z)$ (and a 24th root of it known as the η-function). The η-function occurs in the transformation properties of the Dedekind sums, and the Δ-function occurs in the theory of the moduli of elliptic curves. It also occurs in the limit formula of Kronecker. However, Ramanujan was the first to realize that the coefficients of the q-expansion give rise to an interesting arithmetic sequence.

2 The τ-Function and Partitions

Writing integers as sums of elements of a distinguished subset is a theme that occupied Ramanujan in many of his works. As we said, the τ-function itself arises in the problem of representing an integer as a sum of 24 squares. Ramanujan also gave considerable attention to the partition problem, namely the number of ways of writing a positive integer as a sum of positive integers. If we denote by $p(n)$ the number of such representations of a positive integer n, then we see that the first few values are given by $p(1) = 1$, $p(2) = 2$, $p(3) = 3$, $p(4) = 5$, $p(5) = 7$, $p(6) = 11$ and so on. If we consider the generating function

$$\sum_{n=1}^{\infty} p(n)q^n$$

then it is easily seen to be equal to

$$\prod_{m=1}^{\infty}\big(1-q^m\big)^{-1}.$$

This bears some superficial resemblance to (1). Indeed, such series had been studied classically by Euler, Jacobi and others. However, it was Ramanujan who began to observe arithmetical properties of the coefficients, such as congruences. For example, he observed that

$$p(5n+4) \equiv 0 \pmod 5$$

and

$$p(7n+5) \equiv 0 \pmod 7.$$

He continued this line of thought to the τ-function itself. We know now that such congruences are based on deep aspects of the theory of modular forms.

While the sequence $\{p(n)\}$ and $\{\tau(n)\}$ share properties such as congruence relations, they are, however, very different in terms of their growth properties. Ramanujan's circle method (described in another chapter) shows that $p(n)$ grows exponentially as a function of n, while as we shall see in this chapter, $\tau(n)$ has a polynomial growth in n.

3 Related Generating Functions

Euler had studied the function

$$\prod_{n=1}^{\infty}(1-q^n) \tag{2}$$

and proved that it is equal to

$$\sum_{n\in\mathbb{Z}}(-1)^n q^{(3n^2+n)/2}.$$

Numbers of the form $n(3n+1)/2$ are called *pentagonal*. This and related q-expansion identities can be expressed in terms of the number of partitions of an integer into other integers satisfying various constraints. For example, Euler's identity can be interpreted as giving an expression for the number of partitions of a number into an even number of unequal parts minus the number of partitions of the same number into an odd number of unequal parts.

We can ask for which integers m there is an $n \in \mathbb{Z}$ such that

$$m = \frac{1}{2}n(3n+1)? \tag{3}$$

We need to have $1+24m$ to be a perfect square, say r^2. Moreover, we need $-1 \pm r$ divisible by 6, and in particular, 6 should not divide r. When these conditions are satisfied, we have

$$n = \frac{1}{6}(-1 \pm r).$$

In particular, given m, there are at most two values of n satisfying (3). Thus, when (2) is written as a power series in q, the coefficients are bounded. Moreover, the number of $m \le x$ for which (3) has a solution is $\le \sqrt{2x/3}$. In particular, most of the coefficients are zero, and the series is "lacunary". We can say a little more about which coefficients are nonzero. Indeed, for $1+8m = r^2$ to be satisfied, we need r odd, and

$$r+1 = a, \qquad r-1 = b$$

for some factorization $8m = ab$. This implies that $r = \frac{1}{2}(a+b)$ and $a = b+2$. In particular, both a and b are even, and $r = b+1$ and $8m = b(b+2)$.

Jacobi studied the third power of (2) and proved that

$$\prod_{n=1}^{\infty}\bigl(1-q^n\bigr)^3 = \sum_{n=0}^{\infty}(2n+1)q^{n(n+1)/2}. \tag{4}$$

Numbers of the form $n(n+1)/2$ are called *triangular*. Again, it is seen that this is a lacunary series, in the sense that the set of n for which the coefficient of q^n is nonzero has density zero. Note that we have

$$\sum_{n=1}^{\infty}\tau(n)q^n = q\biggl(\sum_{m\in\mathbb{Z}}(-1)^m q^{(3m^2+m)/2}\biggr)^{24}$$

and

$$\sum_{n=1}^{\infty}\tau(n)q^n = q\biggl(\sum_{m=0}^{\infty}(2m+1)q^{m(m+1)/2}\biggr)^{8}.$$

In particular, one can derive the following formulas for $\tau(n)$:

$$(n-1)\tau(n) = \sum_{1\le|m|\le a_n}\biggl(n-1-\frac{25m}{2}(3m+1)\biggr)\tau\biggl(n-\frac{m}{2}(3m+1)\biggr)$$

where

$$a_n = \frac{1}{6}\bigl(1+(1+24n)^{\frac{1}{2}}\bigr);$$

$$(n-1)\tau(n) = \sum_{1\le m\le b_n}(-1)^{2m+1}(2m+1)\biggl(n-1-\frac{9m}{2}(m+1)\biggr)\tau\biggl(n-\frac{m}{2}(m+1)\biggr)$$

where

$$b_n = \frac{1}{2}\bigl((1+8n)^{\frac{1}{2}}-1\bigr).$$

The first of these is due to Lehmer, and the second to Ramanujan. However, these formulas do not seem to be useful in giving an expression for $\tau(n)$ in terms of elementary functions. They cannot be viewed as 'closed-form' expressions since the range of summation is a function of the argument. However, they do suffice to show that $\tau(n)$ has at most polynomial growth in n.

4 Values of the τ-Function

It is easy to compute the first few values: $\tau(1) = 1$, $\tau(2) = -24$, $\tau(3) = 252$. The following table is copied from [110].

n	$\tau(n)$	n	$\tau(n)$	n	$\tau(n)$
1	1	11	534,612	21	−4,219,488
2	−24	12	−370,944	22	−12,830,688
3	252	13	−577,738	23	18,643,272
4	−1472	14	401,856	24	21,288,960
5	4830	15	1,217,160	25	−25,499,225
6	−6048	16	987,136	26	13,865,712
7	−16744	17	−6,905,934	27	−73,279,080
8	84480	18	2,727,432	28	24,647,168
9	−113,643	19	10,661,420	29	128,406,630
10	−115,920	20	−7,109,760	30	−29,211,840

Looking at this table and others that give more values, many natural questions come to mind. Firstly, we see that the numbers are growing fairly rapidly. However, the growth is not exponential since it was shown by Ramanujan that

$$\left|\tau(n)\right| \ll n^7.$$

He conjectured that

$$\left|\tau(n)\right| \leq d(n)n^{11/2}$$

where $d(n)$ denotes the number of positive divisors of n. This is known as the Ramanujan conjecture (actually Hardy called it the Ramanujan hypothesis), and it is now a theorem as we shall explain in another chapter. The only proof of this relies on "reducing" it to a special case of the Weil conjectures and appealing to the proof of these conjectures by Deligne.

It is a classical result that $d(n) = \mathbf{O}(n^\epsilon)$ for any $\epsilon > 0$, and so a weaker version of the Ramanujan conjecture is that for any $\epsilon > 0$,

$$\tau(n) \ll_\epsilon n^{11/2+\epsilon}$$

where the subscript indicates that the implied constant may depend on ϵ.

We might also notice from the table a fact that Ramanujan stated as a conjecture, namely that the τ-function is multiplicative:

$$\tau(mn) = \tau(m)\tau(n) \tag{5}$$

for positive integers m and n that are relatively prime. Moreover, Ramanujan observed that for a fixed prime p, the values of $\tau(p^m)$ satisfy a second-order recurrence relation: for $m \geq 1$, we have

$$\tau\left(p^{m+1}\right) = \tau(p)\tau\left(p^m\right) - p^{11}\tau\left(p^{m-1}\right). \tag{6}$$

p	$\tau(p)$	p	$\tau(p)$	p	$\tau(p)$
2	−24	31	−52,843,168	73	1,463,791,322
3	252	37	−182,213,314	79	38,116,845,680
5	4830	41	308,120,442	83	−29,335,099,668
7	−16744	43	−17,125,708	89	−24,992,917,110
11	534,612	47	2,687,48,496	97	75,013,568,546
13	−577,738	53	−1,596,055,698	101	81,742,959,102
17	−6,905,934	59	−5,189,203,740	103	−225,755,128,648
19	10,661,420	61	6,956,478,662	107	90,241,258,356
23	18,643,272	67	−15,481,826,884	109	73,482,676,310
29	128,406,630	71	9,791,485,272	113	−85,146,862,638

Both of these properties were proved by Mordell within a year of the publication of Ramanujan's paper. These relations imply that all of the values of the τ-function can be determined once they are known at prime arguments. Above is a table of $\tau(p)$ for primes $p \leq 113$ (the first 30 primes).

Consider again the values of the τ-function on powers of a prime. Denote by α_p and β_p the complex numbers which are roots of the equation

$$T^2 - \tau(p)T + p^{11} = 0.$$

Then $\tau(p) = \alpha_p + \beta_p$ and $\alpha_p\beta_p = p^{11}$. Let us write

$$\alpha_p = p^{11/2}e^{i\theta_p}.$$

Here θ_p is a complex number, and the Ramanujan conjecture is the assertion that in fact θ_p is real. In any case, we deduce that

$$\tau(p) = 2p^{11/2}\cos(\theta_p).$$

This implies that

$$\tau(p)^2 = 4p^{11}\big(1 - \sin^2(\theta_p)\big).$$

We see from this that we cannot have $\sin(\theta_p) = 0$ as $\tau(p)$ is an integer. Moreover, relation (6) shows that

$$\tau\big(p^2\big) = \tau(p)^2 - p^{11} = \alpha_p^2 + \alpha_p\beta_p + \beta_p^2.$$

More generally, by induction, we can show that

$$\tau\big(p^a\big) = \frac{\alpha_p^{a+1} - \beta_p^{a+1}}{\alpha_p - \beta_p}.$$

Equivalently,

$$\tau\left(p^a\right) = p^{11a/2}\frac{\sin(a+1)\theta_p}{\sin\theta_p}. \tag{7}$$

5 Parity of the τ-Function

A few more calculations will show that $\tau(p)$ seems to be even for all primes p. This is in fact true and can be proved as follows. We have the congruence

$$\left(1-q^n\right)^{24} \equiv \left(1+q^{8n}\right)^3 \pmod 2.$$

Now, by a q-series identity of Jacobi, we have

$$\prod_{n=1}^{\infty}\left(1+q^{8n}\right)^3 = \sum_{m=0}^{\infty} q^{4m^2+4m}.$$

Thus, we deduce that

$$\sum_{n=1}^{\infty}\tau(n)q^n \equiv q\sum_{m=0}^{\infty} q^{4m^2+4m} \equiv \sum_{m=0}^{\infty} q^{(2m+1)^2} \pmod 2.$$

In particular, $\tau(n)$ is odd if and only if $n=(2m+1)^2$, in other words, if and only if n is an odd square. In particular, $\tau(p)$ is even for every prime p.

6 Congruences Satisfied by the τ-Function

The result of the previous paragraph is that

$$\tau(p) \equiv 0 \pmod 2$$

for all primes p. There are many other congruence relations discovered by Ramanujan. Here is a partial list:

(1) $\tau(p) \equiv 1 + p^3 \pmod{2^5}$
(2) $\tau(p) \equiv 1 + p \pmod 3$
(3) $\tau(p) \equiv p + p^{10} \pmod{5^2}$
(4) $\tau(p) \equiv p + p^4 \pmod 7$
(5) $\tau(p) \equiv 1 + p^{11} \pmod{691}$

More congruences were discovered later by other authors including Bambah, Chowla, Gandhi, Swinnerton-Dyer and Wilton.

As mentioned briefly earlier, Ramanujan was also the first to find congruences satisfied by the partition function. Some of these are

(1) $p(5m+4) \equiv 0 \pmod 5$
(2) $p(7m+5) \equiv 0 \pmod 7$
(3) $p(11m+6) \equiv 0 \pmod{11}$

In Chap. 7, we will indicate how such congruences can be proved.

7 Vanishing of the τ-Function

Of the many problems that are open with respect to the τ-function, there is a conjecture of Lehmer [111] that asserts that

$$\tau(p) \neq 0$$

where p is a prime. Equivalently,

$$\tau(n) \neq 0$$

for any $n \geq 1$. In fact, we have the following elementary result of Lehmer.

Proposition 7.1 *Let n_0 denote the least value of n for which $\tau(n) = 0$ (if it exists). Then n_0 is prime.*

Proof The multiplicativity of the τ-function (5) shows that n_0 is a prime power, say $n_0 = p^a$. Suppose that $a > 1$ (in other words, $\tau(p) \neq 0$). Then from (7) we deduce that

$$\sin(a+1)\theta_p = 0$$

and so

$$\theta_p = k\pi/(a+1) \tag{8}$$

for some integer k. The number

$$4(\cos\theta_p)^2 = \tau(p)^2/p^{11}$$

is rational. On the other hand, by (8), it is an algebraic integer. Thus, it is in fact an integer, say m, and it is also clear that $m > 0$. Again, by (8), θ_p is real, and so $|\cos\theta_p| \leq 1$ and so $m \leq 4$. Then

$$\tau(p)^2 = mp^{11}$$

and as $\tau(p)$ is an integer, m must be divisible by p. Putting all of these constraints together, we see that $p = m = 2$ or $p = m = 3$. This implies that

$$\tau(2) = \pm 2^6 \quad \text{or} \quad \tau(3) = \pm 3^6.$$

But neither of these can hold as we see from the tables above: $\tau(2) = -24$ and $\tau(3) = 252$. This completes the proof. □

We also have the following curious result.

Proposition 7.2 *Suppose that the set* $\{n : \tau(n) = 0\}$ *has density zero. Then it is, in fact, empty.*

Proof Suppose that $\tau(n) = 0$ for some n. Then by the previous result, the least such n is a prime number, p (say). It follows that if $n = mp$ with $p \nmid m$, then $\tau(n) = 0$. Thus, the set

$$\big\{n : \tau(n) = 0\big\}$$

has density $\geq (p-1)/p^2$, contradicting our hypothesis. □

Using the congruences satisfied by the τ-function, it is possible to show that n_0 (if it exists) must be quite large. Indeed, Lehmer himself observed that if $\tau(p) = 0$, then the congruences imply that

$$p^3 \equiv -1 \pmod{2^5}.$$

Since $p \neq 2$, we also have that

$$p^{16} \equiv 1 \pmod{2^5}.$$

These two congruences together imply that

$$p^2 \equiv 1 \pmod{2^5}$$

from which it follows that

$$p \equiv -1 \pmod{2^5}.$$

Similarly, one finds $p \equiv -1 \pmod{2^5 \cdot 3 \cdot 5^2 \cdot 691}$. This immediately implies that

$$p \geq 2^5 \cdot 3 \cdot 5^2 \cdot 691 - 1 = 1{,}658{,}399.$$

In fact, one can do much better, and it is currently known that $\tau(p) \neq 0$ for $p < 10^{15}$. Of course. Lehmer's conjecture is that it should be nonzero for all p. Recently, this conjecture has been related to the irrationality of coefficients of certain mock modular forms.

There is an even stronger conjecture than that of Lehmer, which has been suggested by Atkin and Serre. For every $\epsilon > 0$, they ask whether

$$\big|\tau(p)\big| \gg_\epsilon p^{\frac{9}{2}-\epsilon}?$$

Nothing is known about this conjecture. In terms of the angle θ_p introduced earlier, the above conjecture is equivalent to the assertion that

$$\left|\theta_p - \frac{\pi}{2}\right| \gg_\epsilon \frac{1}{p^{1+\epsilon}}.$$

If we consider $\tau(n)$ rather than $\tau(p)$, it is possible to prove a kind of lower bound. It is shown in [134] that there is an effectively computable absolute constant $c > 0$ such that for all positive integers n for which $\tau(n)$ is *odd*, we have the lower bound

$$\left|\tau(n)\right| \geq (\log n)^c.$$

The condition on the parity of $\tau(n)$ ensures that it is not divisible by the first power of any prime. Thus, the set of n for which the result applies is the so-called squarefull numbers (that is, numbers for which every prime divisor occurs to at least the second power).

It is also interesting to ask for lower bounds that hold infinitely often. Hardy showed that

$$\tau(n) > n^{11/2}$$

holds infinitely often. The best result in this regard is to due to R. Murty [128], who showed that there is an absolute and effective constant $c > 0$ such that

$$\left|\tau(n)\right| > n^{11/2} \exp\{c \log n / \log\log n\}.$$

This result is essentially best possible since we know that

$$d(n) < \exp\left\{c' \log n / \log\log n\right\}$$

and by Ramanujan's conjecture (Deligne's theorem), we have

$$\left|\tau(n)\right| \leq d(n) n^{11/2}.$$

With recent developments on the Sato–Tate conjecture (see Chaps. 10 and 12), these results are valid for any $c < \log 2$. □

8 Divisibility of $\tau(p)$ by p

The many congruences satisfied by the τ-function all have a fixed modulus and varying argument. A different kind of congruence is the one in the title of this section, namely whether

$$\tau(p) \equiv 0 \pmod{p}. \tag{9}$$

Indeed, this does occur, and the first example is $p = 2$ as we have already seen above. The only primes known to satisfy (9) are $p = 2, 3, 5, 7, 2411, 7758337633$.

(This last prime was discovered quite recently [116].) It is not known whether there are infinitely many such primes. Neither is it known that the complement is infinite! In this regard, the situation is similar to that of Wieferich primes.

Heuristic reasoning would suggest that

$$\#\{p \le x : \tau(p) \equiv 0 \pmod{p}\} \sim \log\log x. \tag{10}$$

However, the function $\log\log x$ grows so slowly that it is computationally difficult to distinguish it from a constant. (For example, $\log\log 10^{15}$ is approximately 3.54.) If (10) were true, we would expect that

$$\sum_{\tau(p)\equiv 0 \pmod{p}} \frac{1}{\log p}$$

converges. However, we do not even know whether

$$\sum_{\tau(p)\equiv 0 \pmod{p}} \frac{1}{p}$$

converges.

Suppose that p is a prime for which (9) holds. Then from the relation

$$\tau\left(p^{a+1}\right) = \tau(p)\tau\left(p^a\right) - p^{11}\tau\left(p^{a-1}\right)$$

we see that for all $a \ge 1$,

$$\tau\left(p^a\right) \equiv 0 \pmod{p^a}.$$

We also note that if $\tau(p) \equiv 0 \pmod{p}$ and $n = pm$ where $(p, m) = 1$, then

$$\left(\tau(n), n\right) \ne 1. \tag{11}$$

If the set of primes for which (9) holds has positive density, then it follows by an elementary sieve argument that the set of n for which (11) holds has density 1. This latter statement can be proved unconditionally. In fact, one knows that

$$\#\{n \le x : \left(\tau(n), n\right) = 1\} \ll x/\log\log\log x.$$

It is interesting to note that if there are infinitely many primes p such that $\tau(p) = 0$, then for any value of k, there exist infinitely many values of n such that $\tau(n) = \tau(n+1) = \tau(n+2) = \cdots = \tau(n+k) = 0$. This is an easy exercise using the Chinese remainder theorem.

9 Lehmer's Conjecture and Harmonic Weak Maass Forms

Finally, we report on some recent work of Bruinier, Ono and Rhoades [27] that opens another line of investigation for Lehmer's conjecture. We give a brief (and slightly technical) explanation of this new development.

After the study of classical modular forms, it is natural to study the space of weakly holomorphic modular forms which are just meromorphic modular forms whose singularities may only occur at the cusps. These spaces are contained in the larger space of harmonic weak Maass forms which need not be holomorphic but are annihilated by a second-order differential operator. At the cusps, we allow for exponential growth of a special type. To be precise, let

$$\Delta_k := -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) + iky\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

Let χ be a Dirichlet character modulo N. A *harmonic weak Maass form of weight k* on $\Gamma_0(N)$ with Nebentypus χ is a smooth function on the upper half-plane satisfying the following three conditions:

(1) $$f\left(\frac{az+b}{cz+d}\right) = \chi(d)(cz+d)^k f(z) \quad \text{for all} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N);$$

(2) $\Delta_k f = 0$;

(3) there is a polynomial $P_f = \sum_{n\le 0} c_f^+(n)q^n \in \mathbb{C}[q^{-1}]$ such that for some fixed $\epsilon > 0$, $f(z) - P_f(z) = O(e^{-\epsilon y})$ as y tends to infinity, with analogous conditions at the other cusps. The polynomial P_f is called the *principal part* of f at the corresponding cusp.

This vector space of harmonic weak Maass forms is denoted $H_k(\Gamma_0(N), \chi)$. One can show that every weight $2-k$ harmonic weak Maass form $f(z)$ has a Fourier expansion at each cusp of the following form:

$$f(z) = \sum_{n \gg -\infty} c_f^+(n)q^n + \sum_{n<0} c_f^-(n)\Gamma\big(k-1, 4\pi|n|y\big)q^n,$$

where $\Gamma(a,x)$ is the incomplete Gamma function given by

$$\Gamma(a,x) = \int_x^\infty t^{a-1}e^{-t}\,dt.$$

The differential operator

$$\xi_w := 2iy^w\overline{\frac{\partial}{\partial \overline{z}}},$$

has the property that

$$\xi_{2-k} : H_{2-k}\big(\Gamma_0(N), \chi\big) \to S_k\big(\Gamma_0(N), \overline{\chi}\big).$$

A straightforward calculation shows that

$$\xi_{2-k}(f) = -(4\pi)^{k-1}\sum_{n=1}^{\infty}\overline{c_f^-(-n)}n^{k-1}q^n.$$

In other words, the coefficients $c_f^-(n)$ are really coefficients of classical cusp forms. Now let $g \in S_k(\Gamma_0(N), \chi)$ be a normalized newform. We say that a harmonic weak Maass form f is *good for* g if the following conditions are satisfied:

(1) The principal part of f at the cusp ∞ belongs to $F_g[q^{-1}]$, where F_g is the number field obtained by adjoining the Fourier coefficients of g to $\mathbb{Q}$;
(2) The principal parts of f at the other cusps of $\Gamma_0(N)$ are constant;
(3) $\xi_{2-k}(f) = g/(g, g)$, where $(\cdot,\cdot)$ denotes the Petersson inner product.

The main theorem of [27] is that if

$$g = \sum_{n=1}^{\infty} c_g(n) q^n \in S_k\big(\Gamma_0(N), \overline{\chi}\big)$$

is a normalized newform and $f \in H_{2-k}(\Gamma_0(N), \chi)$ is good for g, then for any p coprime to N for which $c_g(p) = 0$, we have $c_f^+(n)$ algebraic for any n with $\mathrm{ord}_p(n)$ odd. In other words, the vanishing of the Fourier coefficients of a Hecke eigenform implies the algebraicity of the Fourier coefficients of the corresponding harmonic weak Maass form. The authors in [27] discuss the case for Lehmer's conjecture in this context.

Chapter 3
Ramanujan's Conjecture and ℓ-Adic Representations

Let us recall that Ramanujan's conjecture asserts that for all primes p, we have

$$\left|\tau(p)\right| \leq 2p^{11/2}.$$

The key to understanding this is the modular nature of $\Delta(z)$. There is an action of the modular group

$$\mathrm{SL}_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}, a, b, c, d \in \mathbb{Z},\ ad - bc = 1 \right\}$$

on complex numbers with positive imaginary part, the so-called upper half-plane

$$\mathcal{H} = \left\{ z \in \mathbb{C} : \Im(z) > 0 \right\}.$$

The action is given by fractional linear transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} z = \frac{az + b}{cz + d}.$$

Under this action, Δ has the following transformation property:

$$\Delta(\gamma z) = (cz + d)^{12} \Delta(z).$$

This, and the holomorphy of Δ on $\mathcal{H}$ as well as at the boundary of $\mathcal{H}$, make Δ a modular form. The weight of a modular form is determined by the power of the automorphy factor

$$j(\gamma, z) = (cz + d)$$

that appears in the transformation property. Thus Δ is a modular form of weight 12. The q-expansion

$$\Delta(z) = \sum_{n=1}^{\infty} \tau(n) q^n, \qquad q = e^{2\pi i z}$$

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_3, © Springer India 2013

is then seen as the Fourier expansion of Δ (at the cusp $i\infty$). The fact that the Fourier expansion has no constant term (that is, the coefficient of q^0 is 0) designates Δ as a cusp form.

Petersson generalized Ramanujan's conjecture to other modular forms which are cusp forms. More precisely, he conjectured that for a cusp form of weight k and having a Fourier expansion

$$f(z) = \sum_{n=1}^{\infty} a_n q^n$$

at the cusp $i\infty$, the coefficients satisfy the inequality

$$\left|a(p)\right| \ll_\epsilon p^{(k-1)/2+\epsilon}$$

for any $\epsilon > 0$.

1 The Weil Conjectures

Let X be a smooth projective variety over a finite field $\mathbb{F}$ of q elements. Let us fix an algebraic closure $\overline{\mathbb{F}}$ of $\mathbb{F}$ and for each integer $n \geq 1$, denote by $\mathbb{F}_n$ the unique extension of $\mathbb{F}$ of degree n contained in $\overline{\mathbb{F}}$. The function

$$\zeta(X, T) = \exp\left(\sum_{n=1}^{\infty} \left|X(\mathbb{F}_n)\right| \frac{T^n}{n} \right) \tag{12}$$

is called the zeta function of X. For example, if $X = \mathbb{P}^1$ is the projective line, then $|X(\mathbb{F}_n)| = q^n + 1$, and an easy calculation shows that

$$\zeta\left(\mathbb{P}^1, T\right) = \frac{1}{(1-T)(1-qT)}.$$

More generally, if $X = \mathbb{P}^n$, writing projective space as a stratified union of affine spaces of smaller dimension, we find that

$$\zeta\left(\mathbb{P}^n, T\right) = \frac{1}{(1-T)(1-qT)(1-q^2T)\cdots(1-q^nT)}.$$

Weil made three conjectures about this class of zeta functions. First, he conjectured that it is a rational function of T. More precisely, he conjectured that there are polynomials $P_i(X, T) \in \mathbb{Z}[T]$ for $0 \leq i \leq 2\dim X$ so that

$$\zeta(X, T) = \prod_{i=0}^{2\dim X} P_i(X, T)^{(-1)^{(i+1)}} \tag{13}$$

with $P_0(T) = (1 - T)$ and $P_{2\dim X}(T) = (1 - q^{\dim X}T)$.

He also conjectured that the P_i satisfy a functional equation relating $P_i(X,T)$ to $P_{2\dim X-i}(X,(q^{b_i}T^{a_i}))$ where a_i and b_i are integers determined by the topology of X. Finally, and perhaps most strikingly, he conjectured that the polynomials P_i can be factored as

$$P_i(T)=\prod_{j=1}^{2b_i}(1-\alpha_{i,j}T) \tag{14}$$

and the $\alpha_{i,j}$ are complex numbers satisfying

$$|\alpha_{i,j}|=q^{i/2} \tag{15}$$

for all relevant values of j. This third conjecture is called the Riemann hypothesis for X.

To understand this terminology, we have to rely on another (equivalent) definition of the zeta function. Let us set (see [182]) for a complex number s,

$$\zeta(X,s)=\prod_P\left(1-\frac{1}{N(P)^s}\right)^{-1}$$

where the product is over the closed points P of X, and $N(P)$ is the order of the residue field of P. This is the geometric analogue of the Riemann zeta function. We assume that $\Re(s)$ is sufficiently large so that this product converges. Then, we have the relation

$$\zeta(X,s)=Z\big(X,q^{-s}\big)$$

where q is the cardinality of $\mathbb{F}$. In this case, (15) and (14) show that the zeros of $\zeta(X,s)$ satisfy $\Re(s)=1/2,3/2,\ldots,\dim X-\frac{1}{2}$ and the poles satisfy $\Re(s)=0,1,\ldots,\dim X$. In the sense that (15) allows us to locate the zeros and poles of $\zeta(X,s)$ on certain vertical lines, it can be said to be an analogue of the Riemann hypothesis.

The original motivation of Weil in proposing his conjectures was to address the problem of counting the number of points on varieties over finite fields. If we logarithmically differentiate relations (12), (13) and (14), we deduce that

$$\big|X(\mathbb{F}_n)\big|=\sum_{i=0}^{2\dim X}\sum_{j=1}^{2b_i}(-1)^i\alpha_{i,j}^n$$

and using (15), this becomes

$$\big|X(\mathbb{F}_n)\big|=q^{(\dim X)n}\big\{1+\mathbf{O}_X\big(q^{-n/2}\big)\big\}$$

where the constant implicit in the error term depends on X.

2 The Case of Elliptic Curves

Weil himself proved his conjectures in the case that X is a curve (that is $\dim X = 1$). In this case, the expectation is that

$$Z(X,T) = \frac{P(T)}{(1-T)(1-qT)} \tag{16}$$

where $P(T) \in \mathbb{Z}[T]$, and the degree of $P(T)$ should be $2g$ where g is the genus of the curve X.

The case of elliptic curves is of particular interest. Here we have $g = 1$, and so $P(T)$ is a quadratic polynomial. Understanding this case provides insight into Weil's intuition that there should be a relationship between the topology of X and the number of solutions of various congruences (number of points on X).

Over the complex numbers, an elliptic curve E is a one-dimensional torus

$$E(\mathbb{C}) = \mathbb{C}/L. \tag{17}$$

Here $L \subset \mathbb{C}$ is a lattice, in other words, $L \simeq \mathbb{Z} \oplus \mathbb{Z}$. We see that for any $1 \leq n \in \mathbb{Z}$, we have

$$E[n] = \frac{1}{n}L/L \;\simeq\; \mathbb{Z}/n\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$$

where $E[n]$ denotes the set of points R in $E(\mathbb{C})$ satisfying $nR = 0$. The subgroups $E[n]$ form a directed system under the partial ordering on $\mathbb{Z}$ given by divisibility, with natural morphisms

$$E[n] \longrightarrow E[m]$$

whenever $m|n$ given by

$$R \mapsto \frac{n}{m}R.$$

The inverse limit under these maps is then seen to be

$$L \otimes \hat{\mathbb{Z}} \simeq \prod_{\ell}(L \otimes \mathbb{Z}_\ell).$$

In particular, for each prime ℓ, the ℓ-primary component is denoted $T_\ell(E)$ and is called the Tate module of E. We see that as abelian groups

$$T_\ell(E) \simeq L \otimes \mathbb{Z}_\ell \simeq \mathbb{Z}_\ell \oplus \mathbb{Z}_\ell.$$

For elliptic curves over more general fields F, one no longer has the uniformization given by (17). However, the groups $E[n]$ still make sense, as does $T_\ell(E)$. Moreover, if ℓ is different from the characteristic of F, we still have

$$T_\ell(E) \simeq \mathbb{Z}_\ell \oplus \mathbb{Z}_\ell$$

as abelian groups. But if F is not algebraically closed, there is a new structure that emerges. If we choose an algebraic closure $\overline{F}$ of F, then the $E[n]$ and $T_\ell(E)$ are not only abelian groups, but modules for the natural action of $\mathrm{Gal}(\overline{F}/F)$. Thus, we have a representation

$$\mathrm{Gal}(\overline{F}/F) \longrightarrow \mathrm{Aut}_{\mathbb{Z}_\ell}\big(T_\ell(E)\big) \simeq GL_2(\mathbb{Z}_\ell).$$

In particular, if $F = \mathbb{F}$ is a finite field of q elements, we have the Frobenius automorphism $\sigma \in \mathrm{Gal}(\overline{F}/F)$ which is given by

$$\sigma(x) = x^q.$$

For $(\ell, q) = 1$, we thus get a matrix in $GL_2(\mathbb{Z}_\ell)$. If we denote by $P_\ell(T)$ the characteristic polynomial of this matrix, it is a fact that

$$P_\ell(T) = P(T), \tag{18}$$

where $P(T)$ is given by (16). In particular, $P_\ell(T)$ has coefficients in $\mathbb{Z}$ (instead of just $\mathbb{Z}_\ell$) and is independent of ℓ.

If we now consider an elliptic curve E over a number field F, the above constructions still make sense, and we can speak of the Tate module $T_\ell(E)$. We again have the family of Galois representations

$$\mathrm{Gal}(\overline{F}/F) \longrightarrow \mathrm{Aut}_{\mathbb{Z}_\ell}\big(T_\ell(E)\big) \simeq GL_2(\mathbb{Z}_\ell) \tag{19}$$

giving the action of the Galois group on points on E of ℓ-power order.

If v is a prime of F, we may consider the reduction E_v of E modulo v. Naively, this just means that we reduce the coefficients of an equation defining E modulo v. Apart from a finite number of primes v of F, we have that E_v is again an elliptic curve, and it is defined over the residue field $\mathbb{F}$, which is a field of cardinality equal to the norm $\mathbb{N}v$ of v.

As in the previous paragraph, we may consider (now for ℓ not dividing $\mathbb{N}v$), the Tate module $T_\ell(E_v)$. On the other hand, we may consider the original $T_\ell(E)$ but now viewing it as a module restricted to just the decomposition group at v. It is a theorem that under the assumption that E_v is an elliptic curve, the inertia subgroup at v acts trivially on $T_\ell(E)$, and thus it may also be viewed as a module for $\mathrm{Gal}(\overline{\mathbb{F}}/\mathbb{F})$. (The converse of this theorem is also true. Thus, if the inertia subgroup at v acts trivially on $T_\ell(E)$, then the reduction E_v is again an elliptic curve.)

Now we have two modules for $\mathrm{Gal}(\overline{\mathbb{F}}/\mathbb{F})$, namely $T_\ell(E)$ and $T_\ell(E_v)$. It turns out that they are isomorphic. In particular, we may speak of the characteristic polynomial of Frobenius at v acting on $T_\ell(E)$, and it is the same $P_\ell(T) = P(T)$ that we found in (18).

Considering again the global Galois representation (19), we now have a family of ℓ-adic representations which have the property that for any prime v of the number field F at which the inertia subgroup acts trivially, the characteristic polynomial of Frobenius at v makes sense, and if we denote it by $P_\ell(T, v)$, then it is an element of $\mathbb{Z}[T]$ and is independent of ℓ.

Moreover, $P_\ell(T, v)$ is the numerator of the zeta function of an elliptic curve E_v over a finite field. It is thus of the form

$$P_\ell(T, v) = 1 - a_v T + (Nv)T^2$$

where $a_v \in \mathbb{Z}$, and applying the Riemann hypothesis for E_v, it follows that

$$|a_v| \leq 2(Nv)^{\frac{1}{2}}. \tag{20}$$

We may also use this information to construct a formal product

$$\prod_v \left(1 - a_v(Nv)^{-s} + (Nv)^{1-2s}\right)^{-1}.$$

The product only ranges over primes v for which E_v is an elliptic curve. Using estimate (20), we see that the formal product actually converges absolutely for $Re(s) > 3/2$. Up to finitely many factors (namely those at primes v for which E_v is not an elliptic curve), this function is called the L-function associated to E and denoted $L(E, s)$. By analogy with the definition in the first section, the zeta function of E is given by

$$\zeta(E, s) = \zeta(s)\zeta(1 - s)/L(E, s).$$

3 ℓ-Adic Representations

More generally, let us consider a number field F and a family M_ℓ of finite-dimensional $\mathbb{Z}_\ell$ modules which carry an action of $\mathrm{Gal}(\overline{F}/F)$. We say that the representation (or equivalently, the module M_ℓ)

$$\mathrm{Gal}(\overline{F}/F) \longrightarrow \mathrm{Aut}_{\mathbb{Z}_\ell}(M_\ell)$$

is unramified at a prime v of F if the inertia subgroup at v acts trivially. We say that the family $\{M_\ell\}$ is unramified at v if every element M_ℓ of the family is unramified at v. Finally, we say that $\{M_\ell\}$ is a compatible family of ℓ-adic representations if there is a finite set S of primes of F such that for all primes $v \notin S$, the family is unramified at v and for any prime ℓ that is relatively prime to

$$Nv \prod_{w \in S} Nw,$$

the characteristic polynomial $P_\ell(T, v)$ is independent of ℓ and has coefficients in $\mathbb{Z}$.

There are many such families that arise from algebraic geometry. The example of elliptic curves was seen in the previous section. Now the question is whether there is such a family with $F = \mathbb{Q}$ and unramified at all primes, and with the property that for $p \neq \ell$,

$$P_\ell(T, p) = 1 - \tau(p)T + p^{11}T^2.$$

If there were such a representation, then there would be the hope that the Ramanujan conjecture could be put into a context where the Weil conjectures would apply. In that case, it might even be possible to deduce the Ramanujan conjecture from the Weil conjectures.

4 Elliptic Curves and Modular Forms

An example may illustrate what we can expect. Writing

$$q\prod_{n=1}^{\infty}(1-q^n)^{24}=q\prod_{n=1}^{\infty}(1-q^n)^2\prod_{n=1}^{\infty}(1-q^n)^{22}$$

we see that

$$\sum_{n=1}^{\infty}\tau(n)q^n\equiv q\prod_{n=1}^{\infty}(1-q^n)^2\prod_{n=1}^{\infty}(1-q^{11n})^2 \pmod{11}.$$

Using the notation of the chapter on the Ramanujan τ function, the right-hand side of the congruence can be written as $\eta(z)^2\eta(11z)^2$ (where as usual, we are writing $q=e^{2\pi iz}$). Expanding the product, and writing

$$\eta(z)^2\eta(11z)^2=\sum_{n=1}^{\infty}a_nq^n$$

we have the following table of values:

n	a_n	n	a_n	n	a_n
1	1	6	2	11	−1
2	−2	7	−2	12	−2
3	−1	8	0	13	4
4	2	9	−2	14	4
5	1	10	−2	15	−1

The function $f(z)=\eta(z)^2\eta(11z)^2$ is an example of a modular form. It is a holomorphic function defined for complex numbers z with the property that $\Im(z)>0$. Moreover, it satisfies a functional equation with respect to certain fractional linear transformation. Namely, if we denote

$$\Gamma_0(N)=\left\{\begin{pmatrix}a & b\\ c & d\end{pmatrix}, a,b,c,d\in\mathbb{Z},\ ad-bc=1, N|c\right\}$$

then

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 f(z).$$

Since

$$\frac{d}{dz}\left(\frac{az+b}{cz+d}\right) = (cz+d)^{-2}$$

it follows that $f(z)\,dz$ is invariant under the action of $\Gamma_0(N)$.

Now we define a map

$$\Phi_f : \Gamma_0(N) \longrightarrow \mathbb{C}$$

as follows. Fix a point $z_0 \in \mathbb{C}$ with $\Im(z_0) > 0$. Set

$$\Phi_f(\gamma) = \int_{z_0}^{\gamma z_0} f(z)\,dz.$$

The integral on the right does not depend on z_0. Indeed, we have

$$\int_{z_1}^{\gamma z_1} = \int_{z_1}^{z_0} + \int_{z_0}^{\gamma z_0} + \int_{\gamma z_0}^{\gamma z_1}.$$

By a change of variables and the invariance of $f(z)\,dz$ under the action of γ, we see that

$$\int_{\gamma z_0}^{\gamma z_1} f(z)\,dz = \int_{z_0}^{z_1} f(z)\,dz = -\int_{z_1}^{z_0} f(z)\,dz.$$

Thus,

$$\int_{z_1}^{\gamma z_1} f(z)\,dz = \int_{z_0}^{\gamma z_0} f(z)\,dz.$$

Note also that as the upper half-plane is simply connected, the value of the integral does not depend on the path chosen from z_0 to γz_0.

The map Φ_f is in fact a homomorphism since for $\gamma_1, \gamma_2 \in \Gamma_0(N)$, we have

$$\Phi_f(\gamma_1\gamma_2) = \int_{z_0}^{\gamma_2 z_0} f(z)\,dz + \int_{\gamma_2 z_0}^{\gamma_1\gamma_2 z_0} f(z)\,dz.$$

Again by a change of variables and using the invariance of $f(z)\,dz$ under γ_2, we see that this is equal to

$$\Phi_f(\gamma_1\gamma_2) = \Phi_f(\gamma_1) + \Phi_f(\gamma_2).$$

The elements in $\Gamma_0(N)$ can be classified into three types depending on their trace. Thus an element

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is called elliptic if $|\operatorname{trace}(\gamma)| = |a + d| < 2$, hyperbolic if $|\operatorname{trace}(\gamma)| > 2$ and parabolic of $|\operatorname{trace}(\gamma)| = 2$.

Proposition 4.1 *If $\gamma \in \Gamma_0(N)$ is elliptic or parabolic, then $\Phi_f(\gamma) = 0$.*

Proof If γ is elliptic, then it has finite order. Since $\Phi_f(1) = 0$, and Φ_f is a homomorphism, it follows that $\Phi_f(\gamma) = 0$ for any element of finite order. If γ is parabolic, then $a + d = \pm 2$. Note that $-I \in \Gamma_0(N)$. Moreover, as it is of order 2, by the comment above, $\Phi_f(-I) = 0$. Hence, multiplying γ by $-I$ if necessary, we may suppose that $a + d = 2$. Suppose first that $c = 0$. Then

$$\gamma = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}.$$

Then

$$\Phi_f(\gamma) = \int_{z_0}^{z_0+r} f(z)\,dz.$$

As f is a cusp form, its constant term at the cusp $i\infty$ vanishes, and so

$$\int_{z_0}^{z_0+1} f(z)\,dz = 0.$$

Iterating this, we deduce that $\Phi_f(\gamma) = 0$. Now suppose that $c \neq 0$. As γ is parabolic, it fixes the rational point $(a - 1)/c$. Choose a matrix $\beta \in \mathrm{SL}_2(\mathbb{Z})$ such that β maps this point to $i\infty$. Then

$$\gamma_1 = \beta\gamma\beta^{-1}$$

fixes $i\infty$. We have

$$\int_{z_0}^{\gamma z_0} f(z)\,dz = \int_{z_0}^{\beta^{-1}\gamma_1\beta z_0} f(z)\,dz.$$

The right-hand side is equal to

$$\int_{\beta z_0}^{\gamma_1\beta z_0} \big(f|\beta^{-1}\big)(z)\,dz = \int_{\beta z_0}^{\beta z_0+r} \big(f|\beta^{-1}\big)(z)\,dz.$$

And this is zero as f is a cusp form and the constant term of $f|\beta^{-1}$ vanishes. This proves the result.

Let us set

$$\Lambda_f = \big\{\Phi_f(\gamma) : \gamma \in \Gamma_0(N)\big\}.$$

It is a theorem of Eichler and Shimura that if f has integer Fourier coefficients, then Λ_f is a lattice in $\mathbb{C}$. (In general, if the Fourier coefficients lie in a number field, we would consider a higher-dimensional version of Λ_f which would give rise to a lattice in $\mathbb{C}^g$ for some $g > 1$.)

Using this lattice, one may then consider the torus $\mathbb{C}/\Lambda_f$ which is an elliptic curve. At first, it is a curve defined over the complex numbers. However, the theory of Eichler and Shimura shows that it in fact has a model defined over the rationals $\mathbb{Q}$.

As an example, consider $\Gamma_0(11)$. It is generated by the three elements

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad V_4 = \begin{pmatrix} 8 & 1 \\ -33 & -4 \end{pmatrix}, \qquad V_6 = \begin{pmatrix} 9 & 1 \\ -55 & -6 \end{pmatrix}.$$

Since T is parabolic, $\Phi_f(T) = 0$. If we set $\omega_1 = \Phi_f(V_4)$ and $\omega_2 = \Phi_f(V_6)$, then Λ_f is the lattice spanned by ω_1 and ω_2. In fact, the corresponding elliptic curve is given by the equation

$$y^2 + y = x^3 - x^2 - 10x - 20$$

and is denoted $X_0(11)$.

As discussed earlier, we have a compatible family of ℓ-adic Galois representations

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathrm{Aut}_{\mathbb{Z}_\ell}\big(T_\ell\big(X_0(11)\big)\big) \simeq GL_2(\mathbb{Z}_\ell)$$

which are unramified outside 11. Moreover, recall that we have set

$$f(z) = \sum_{n=1}^{\infty} a_n q^n.$$

Then for any prime $p \neq 11$, the characteristic polynomial of the Frobenius automorphism at p is

$$1 - a_p T + pT^2.$$

If we "reduce" the 11-adic representation modulo 11, that is, if we compose it with the natural map

$$GL_2(\mathbb{Z}_{11}) \longrightarrow GL_2(\mathbb{Z}/11\mathbb{Z})$$

then the characteristic polynomial of the Frobenius automorphism at p is

$$1 - \big(a_p \pmod{11}\big)T + \big(p \pmod{11}\big)T^2.$$

Now by the congruence given at the start of this section, it follows that this is the same as

$$1 - \big(\tau(p) \pmod{11}\big)T + \big(p \pmod{11}\big)T^2.$$

It is a theorem of Shimura that the map

$$\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow GL_2(\mathbb{Z}/11\mathbb{Z})$$

is surjective. It follows from an application of the Chebotarev density theorem that for any residue classes $a, b \in \mathbb{Z}/11\mathbb{Z}$ with $b \neq 0$, we can find infinitely many primes

(in fact a positive density set of primes) p for which

$$\tau(p) \equiv a \pmod{11}$$

and

$$p \equiv b \pmod{11}.$$

Thus, there cannot exist any special congruences for the τ function modulo 11.

We see from this that the theory of ℓ-adic representations gives us a tool to "understand" congruences satisfied by the τ function. However, the construction described in this section only gives us a "mod ℓ"-representation and that too, only for one value of ℓ, namely $\ell = 11$. The question now becomes where one should look for the general family of representations.

Since the motivation in this chapter is Ramanujan's conjecture on the size of $\tau(p)$, we have to look for these representations in algebraic geometry. □

5 Geometric Realization of Modular Forms of Higher Weight

We begin by giving a brief description of how modular forms of higher weight occur in the cohomology of certain algebraic varieties. Let Γ denote a subgroup of finite index of $SL_2(\mathbb{Z})$. It acts by fractional linear transformations on the upper half-plane:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{az+b}{cz+d}.$$

Consider now the product $\mathcal{H} \times \mathbb{C}$ of the upper half-plane and the complex line. There is an action of $\Gamma \rtimes \mathbb{Z}^2$ on this product by

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, (m,n) \right) \cdot (z,w) = \left(\frac{az+b}{cz+d}, (cz+d)^{-1}(w+mz+n) \right).$$

We consider the quotient $(\Gamma \rtimes \mathbb{Z}^2)\backslash(\mathcal{H} \times \mathbb{C})$. This is in fact an open algebraic surface, and there is a compactification B_Γ of it called the 'universal elliptic curve'. The terminology comes from the fact that it is a fibre variety over the modular curve. The modular curve itself parameterizes elliptic curves with additional structure, and the fibre at a point of the modular curve 'is' the elliptic curve corresponding to that point.

Now, let us consider holomorphic 2-forms on B_Γ. Let f be a holomorphic modular form of weight 3 for Γ and consider the 2-form $f(z)\,dz \wedge dw$ on $\mathcal{H} \times \mathbb{C}$. It is invariant under the action of $\Gamma \rtimes \mathbb{Z}^2$. Indeed, an element

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}, (m,n) \right)$$

acts on f by

$$f\left(\frac{az+b}{cz+d}\right)d\left(\frac{az+b}{cz+d}\right)\wedge d\left(\frac{w+mz+n}{cz+d}\right)$$

which is

$$(cz+d)^3 f(z)(cz+d)^{-2}\,dz\wedge(cz+d)^{-1}\,dw = f(z)\,dz\wedge dw.$$

Because of this invariance, $f(z)\,dz\wedge dw$ gives rise to a holomorphic 2-form on the quotient and on the surface B_Γ. In fact, it is true that all holomorphic 2-forms on B_Γ can be obtained in this way in the sense that there is an isomorphism of the space of cusp forms on Γ of weight 3 and holomorphic 2-forms on B_Γ.

One can go further and analyze the entire Hodge decomposition of $H^2(B_\Gamma,\mathbb{C})$. We have

$$H^2(B_\Gamma,\mathbb{C}) = H^0\big(B_\Gamma,\Omega^2\big)\oplus\overline{H^0\big(B_\Gamma,\Omega^2\big)}\oplus H^{1,1}.$$

The above discussion identifies the first two pieces. The third piece, $H^{1,1}$ turns out to consist entirely of rational (and hence, by Lefschetz, algebraic) classes. This is a result of Shioda [191].

In general, in the construction of B_Γ, if we replace $\mathbb{C}$ with $\mathbb{C}^{k-2}$, one gets what is known as a Sato variety $B_\Gamma^{(k)}$. This is a variety of dimension $k-1$, fibred over the modular curve as before. The fibre at a given point of the modular curve is the $(k-2)$nd power of 'the' elliptic curve corresponding to that point. Again, one can show that the space of holomorphic $(k-1)$ forms on $B_\Gamma^{(k)}$ is isomorphic to the complex vector space of weight k cusp forms for Γ. In general, we do not know much about the field of definition of $B_\Gamma^{(k)}$. However, if Γ is a congruence subgroup, then we know from the general theory of Shimura and others that it is defined over a number field.

Now let us restrict our attention to Γ a congruence subgroup, say $\Gamma_0(N)$. All of the discussion above also has an ℓ-adic analogue. In this case, the complex cohomology referred to above is replaced with ℓ-adic cohomology. This theory is a vast generalization of the theory of the Tate module discussed earlier. In fact, the first ℓ-adic homology of an elliptic curve is its ℓ-adic Tate module.

In general, one considers the ℓ-adic cohomology of $B_\Gamma^{(k)}$ (or rather, a model of this variety defined over the rationals), and one finds that $H_\ell^{k-1}(B_\Gamma^{(k)})$ contains a subspace S_ℓ (say) of dimension $2\dim S_k(\Gamma)$. Moreover, $H_\ell^{k-1}(B_\Gamma^{(k)})$ has two important actions on it. Firstly, it is a module for the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and secondly, it is a module for the Hecke algebra, an algebra of operators referred to in an earlier chapter. The subspace S_ℓ is invariant under both of these actions. Diagonalizing the Hecke action, we get a decomposition

$$S_\ell = \oplus_f S_\ell(f)$$

into submodules, where each $S_\ell(f)$ is a rank two module over $K_f[\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})]$, and K_f is a certain number field (the field generated by the Fourier coefficients of f).

These are the Galois representations attached to modular forms of higher weight whose existence had been conjectured by Serre and constructed by Deligne [38].

Moreover, the Galois action is unramified outside primes dividing ℓN, and if $p \nmid \ell N$, then the characteristic polynomial of the Frobenius conjugacy class at p is

$$\prod \left(1 - a_f(p)^\sigma T + p^{k-1} T^2\right)$$

the product ranging over embeddings $\sigma : K_f \hookrightarrow \overline{\mathbb{Q}_\ell}$. If we factor the quadratic polynomials above, we get complex numbers $\pi_f(p)^\sigma$ such that

$$a_f(p)^\sigma = \pi_f(p)^\sigma + \overline{\pi_f(p)^\sigma}$$

and

$$\pi_f(p)^\sigma \overline{\pi_f(p)^\sigma} = p^{k-1}.$$

It is the reduction of the variety $B_\Gamma^{(k)}$ that one applies the Weil conjectures to. By the above remark about the characteristic polynomial of Frobenius, we see that when $B_\Gamma^{(k)}$ is seen as a variety over $\mathbb{F}_p$, in the action of Frobenius on $H_\ell^{k-1}(B_\Gamma^{(k)} \times \mathbb{F}_p)$, amongst the eigenvalues of Frobenius will be the $\{\pi_f(p)\}$ and by the Weil conjectures, these numbers are all of complex absolute value $p^{(k-1)/2}$. This is, in brief, how the Ramanujan conjecture is deduced from the Weil conjectures. The reader may find the survey article [91] also helpful.

Chapter 4
The Ramanujan Conjecture from $GL(2)$ to $GL(n)$

1 The Ramanujan Conjectures

In retrospect, one sees a progression of ideas from Ramanujan's work on the τ-function to Hecke's theory of modular forms, and then moving to the representation theoretic perspective, we see this progression linking itself to theories of Harish-Chandra and Langlands. Ramanujan's conjecture now sits as a special case of a more comprehensive conjecture in the Langlands program. The survey article [141] provides an excellent introduction to this chain of ideas, and we recommend this to the reader. But the origins of the chain go back to the 1916 paper of Ramanujan which acted as a catalyst for this development.

In his epic paper of 1916, Ramanujan [162] considered the function

$$\Delta(z) = q\prod_{n=1}^{\infty}\left(1-q^n\right)^{24}, \quad q = e^{2\pi iz}.$$

As we have seen, expanding the right-hand side as a power series in q defines the celebrated Ramanujan τ-function:

$$q\prod_{n=1}^{\infty}\left(1-q^n\right)^{24} = \sum_{n=1}^{\infty}\tau(n)q^n.$$

In his paper, Ramanujan made three conjectures concerning $\tau(n)$:

(1) $\tau(mn) = \tau(m)\tau(n)$, for $(m,n) = 1$,
(2) for p prime and $a \geq 1$, $\tau(p^{a+1}) = \tau(p)\tau(p^a) - p^{11}\tau(p^{a-1})$,
(3) $|\tau(p)| \leq 2p^{11/2}$.

The first two conjectures were proved a year later by Mordell [126]. His proof was reproduced by Hardy in his 1936 lectures on Ramanujan's work delivered at Harvard University. Almost twenty years after Ramanujan's paper, Hecke [77] published the conceptual framework from which (1) and (2) emerge as special cases of a larger theory, now called Hecke theory. Despite heroic efforts to settle conjecture (3) by such luminaries as Hardy, Kloosterman, Rankin and Selberg, we had to

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_4, © Springer India 2013

wait until 1974 when Deligne proved it as a consequence of his proof of the Weil conjectures.

We outline the proofs of these conjectures. As mentioned in the introduction to Chap. 3, the essential property of $\Delta(z)$ is that it is a modular form of weight 12 for the full modular group $SL_2(\mathbb{Z})$, which is the group of 2×2 matrices with integer entries and determinant 1. This means that for z in the upper half-plane $\mathfrak{h}$, we have

$$\Delta\left(\frac{az+b}{cz+d}\right) = (cz+d)^{12}\Delta(z) \quad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}). \tag{21}$$

The set of holomorphic functions $f : \mathfrak{h} \to \mathbb{C}$ vanishing at infinity and satisfying the modular relation

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z) \quad \forall \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z})$$

forms a finite-dimensional vector space S_k over $\mathbb{C}$. Since $-I \in SL_2(\mathbb{Z})$, the modular relation shows that $S_k = 0$ for k odd. For k even, one can show without too much difficulty [183] that $\dim S_2 = 0$ and $\dim S_k = [k/12]$ if $k \not\equiv 2 \pmod{12}$. Thus, S_{12} is one-dimensional and spanned by Δ. In fact, 12 is the first value of k for which $S_k \neq 0$.

Hecke discovered a family of linear transformations T_n of the finite-dimensional vector space S_k. These are now called Hecke operators, though they are nascent in Mordell's work. To describe these, it is convenient to introduce the following notation.

Let $GL_2^+(\mathbb{R})$ denote the group of 2×2 matrices with real entries and positive determinant. It is easy to see that $GL_2^+(\mathbb{R})$ acts on $\mathfrak{h}$ via the usual fractional linear transformations:

$$\gamma z := \frac{az+b}{cz+d}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2^+(\mathbb{R}).$$

Given $f \in S_k$, we define the "slash operator" via

$$(f|\gamma)(z) := (\det \gamma)^{k/2}(cz+d)^{-k} f(\gamma z).$$

Thus, $f \in S_k$ if and only if $f|\gamma = f$ for all $\gamma \in SL_2(\mathbb{Z})$.

Hecke observed that $SL_2(\mathbb{Z})$ acts on the set M_n of matrices with integer entries and determinant n. Indeed, with $\Gamma = SL_2(\mathbb{Z})$, we have

$$\Gamma M_n = M_n \Gamma = M_n.$$

Consequently, we can decompose M_n as a disjoint union of Γ-orbits. In fact, a simple application of the division algorithm shows that

$$M_n = \coprod_{ad=n, d>0} \coprod_{b=0}^{d-1} \Gamma \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.$$

Thus, the finite set of matrices

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}, \quad ad = n, d > 0, \quad 0 \le b \le d-1,$$

comprises a complete set of orbit representatives for this action. Hecke defines

$$T_n(f) = n^{k/2-1} \sum_{ad=n, d>0} \sum_{b=0}^{d-1} f \Big| \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}. \tag{22}$$

Since $\Gamma M_n = M_n \Gamma = M_n$, it is easily seen that for any $\gamma \in \Gamma$, we have

$$T_n(f)|\gamma = T_n(f).$$

Noting that $T_n(f)$ also vanishes at infinity whenever f does, shows that $T_n(f) \in S_k$. In other words, for each natural number n, we have a linear transformation $T_n : S_k \to S_k$. These transformations are called Hecke operators. Since

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL_2(\mathbb{Z}),$$

every $f \in S_k$ has a Fourier expansion

$$f(z) = \sum_{m=1}^{\infty} \lambda_f(m) e^{2\pi i m z}.$$

The expansion starts from 1 since the function is holomorphic and vanishes at infinity. From (22) one can routinely compute the Fourier expansion of $T_n(f)$. Indeed,

$$f \Big| \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} (z) = n^{k/2} d^{-k} \sum_{m=1}^{\infty} \lambda_f(m) e^{2\pi i m(az+b)/d}.$$

Inserting this into the inner sum of (22) and applying the elementary fact

$$\sum_{b=0}^{d-1} e^{2\pi i m b/d} = 0,$$

unless $d|m$, in which case it is d, we obtain that

$$T_n(f) = n^{k-1} \sum_{ad=n, d>0} d^{1-k} \left(\sum_{m=1, d|m}^{\infty} \lambda_f(m) e^{2\pi i m a z/d} \right).$$

Writing $m = dr$, we can re-write the inner sum as

$$\sum_{r=1}^{\infty} \lambda_f(dr) e^{2\pi i r a z}.$$

Collecting the terms with $ra = m$, we see that

$$T_n(f) = n^{k-1} \sum_{m=1}^{\infty} \Bigg(\sum_{ad=n, ar=m} \lambda_f(dr) d^{k-1} \Bigg) e^{2\pi i m z}.$$

The coefficient can be re-written as

$$\sum_{a|n,m} \lambda_f\big(nm/a^2\big)(n/a)^{1-k}$$

so that

$$T_n(f) = \sum_{m=1}^{\infty} \Bigg(\sum_{a|n,m} \lambda_f\big(nm/a^2\big) a^{k-1} \Bigg) e^{2\pi i m z}. \tag{23}$$

In the case of $f = \Delta$, we see immediately that $T_n(\Delta) = c_n \Delta$ since S_{12} is one-dimensional. The constant c_n is computed easily by comparing the first Fourier coefficient of both sides. We obtain

$$T_n(\Delta) = \tau(n)\Delta. \tag{24}$$

If $(m, n) = 1$, then computing the mth Fourier coefficient of the left-hand side of (24) using (23) shows that it is $\tau(mn)$. Comparing this with the mth Fourier coefficient of the right-hand side of (24), we obtain $\tau(mn) = \tau(m)\tau(n)$ whenever $(m, n) = 1$, which is Ramanujan's first conjecture.

The second conjecture also follows directly from (23). Indeed, for p prime,

$$T_p(\Delta) = \tau(p)\Delta$$

so that by comparing the p^ath Fourier coefficient of both sides leads to

$$\tau\big(p^{a+1}\big) + p^{11}\tau\big(p^{a-1}\big) = \tau(p)\tau\big(p^a\big),$$

which is Ramanujan's second conjecture.

The significance of Ramanujan's second conjecture is that the power series

$$\sum_{a=0}^{\infty} \tau\big(p^a\big) x^a$$

is a rational function and can be written as

$$\big(1 - \tau(p)x + p^{11}x^2\big)^{-1}.$$

With regard to the third conjecture, there is a simple argument that leads to the estimate $\tau(p) = O(p^6)$. Though this was first proved by Hardy, his proof was more elaborate. We need only make two observations. The first is that a fundamental domain D for the action of $SL_2(\mathbb{Z})$ on the upper half plane can be taken to be the

standard one, namely, $|z| \geq 1, |\Re(z)| \leq 1/2$. In other words, every element in the upper half-plane is $SL_2(\mathbb{Z})$-equivalent to some element of D, and no two interior elements of D are $SL_2(\mathbb{Z})$-equivalent. The second observation is that

$$y^{12}\big|\Delta(z)\big|^2, \quad z = x + iy$$

is invariant under the action of $SL_2(\mathbb{Z})$ as is easily checked using the modular transformation for Δ. From the Fourier expansion of $\Delta(z)$ we see that it has exponential decay at infinity, and therefore it is bounded (say by K) in the fundamental domain D. Since

$$\Delta(x + iy) = \sum_{n=1}^{\infty} \tau(n) e^{-2\pi ny} e^{2\pi inx}$$

we can calculate the nth Fourier coefficient by

$$\tau(n)e^{-2\pi ny} = \int_{-1/2}^{1/2} \Delta(x + iy)e^{-2\pi inx}\, dx.$$

Since $|\Delta(z)| \leq K/y^6$, we deduce

$$\big|\tau(n)\big|e^{-2\pi ny} \leq K/y^6.$$

Choosing $y = 1/n$ gives the estimate $\tau(n) = O(n^6)$.

With Ramanujan, we can consider the Dirichlet series

$$L_\Delta(s) := \sum_{n=1}^{\infty} \frac{\tau(n)}{n^s}$$

which by our modest estimate of $\tau(n)$ converges absolutely for $\Re(s) > 7$. By Ramanujan's conjecture (1) and (2), we can write this as an Euler product:

$$\sum_{n=1}^{\infty} \frac{\tau(n)}{n^s} = \prod_p \left(1 - \frac{\tau(p)}{p^s} + \frac{1}{p^{2s-11}}\right)^{-1}.$$

The modular relation (21) implies that

$$\Delta(-1/iy) = y^{12}\Delta(iy)$$

and this can be used to derive an analytic continuation and functional equation for $L_\Delta(s)$. Apparently, this was first done by Wilton in 1928. This is a major theme of Hecke theory developed systematically, again, much later in the 1930s.

Seated in the comfort of the modern era, it is easy for us to be surprised that some obvious developments did not take place earlier or with greater rapidity. We must however understand that much of this work seems to be against the background of two world wars and there was no one giving us the "bigger picture". In his book,

Lang [106] wrote, "Partly because of Hitler and the war, which almost annihilated the German school of mathematics, and partly because of the great success of certain algebraic methods of Artin, Hasse, and Deuring, modular forms and functions were to a large extent ignored by most mathematicians for about thirty years after the 1930s. Eichler, Maass, Petersson and Rankin were the main exceptions."

This was the case with even Hecke theory. Much more is true about the Hecke operators. A straightforward calculation using (23) shows that $T_n T_m = T_m T_n$ for all m, n. It was only in 1939 that Hecke's student, Hans Petersson introduced an inner product on S_k by defining

$$(f, g) = \int_D y^k f(z)\overline{g(z)}\,\frac{dx\,dy}{y^2}.$$

He was able to show that

$$\big(T_n(f), g\big) = \big(f, T_n(g)\big) \tag{25}$$

so that the T_ns are a commutative family of Hermitian operators. Already in the proof of (25), it was clear that one needed to move to "higher levels" and not just confine oneself to $SL_2(\mathbb{Z})$. Petersson's theorem allowed him to prove that S_k has a basis of eigenfunctions for all of the Hecke operators T_n since (by a familiar theorem of linear algebra) the T_ns could be simultaneously diagonalized.

The theory generalizing the first two conjectures of Ramanujan to other congruence subgroups of $SL_2(\mathbb{Z})$ was slow in maturing. It was only in 1970, when Atkin and Lehner [12] published their fundamental paper that the first part of the theory reached completion.

As for the third conjecture, this had several strands of development. The first strand was largely a development of analytic methods of attack beginning with Kloosterman [98], who, in 1927, applied the circle method and at the same time devised a new way of handling the major arcs, thereby discovering the celebrated Kloosterman sum (about which more will be said below). Kloosterman was able to show that $\tau(p) = O(p^{6-1/8})$ by his method.

In the early 1930s, Salie and Davenport (independently) refined earlier methods and were able to show $\tau(p) = O(p^{6-1/6})$. In the late 1930s, Hardy's doctoral student, R.A. Rankin [169] developed a new method using Eisenstein series that enabled one to get better estimates on $\tau(n)$. This new method came to be known as Rankin's method (or sometimes, the Rankin–Selberg method, since Selberg arrived at it independently). This method (which we will discuss later in this chapter) gives $\tau(p) = O(p^{6-1/5})$.

Yet another analytic method using the theory of Poincaré series was developed by Selberg. This method allows one to get an explicit formula for $\tau(n)$ that involves Kloosterman sums and Bessel functions. Selberg [180] showed that this explicit formula leads to the improved estimate $O(p^{6-1/4})$.

Selberg's work led to many further and fundamental developments, most notable of which is the Selberg trace formula. In a special case, this becomes the Eichler–Selberg trace formula which gives a formula for the trace of the nth Hecke operator

acting on the space of cusp forms in terms of class numbers of imaginary quadratic fields. In particular, it relates the Ramanujan τ-function to such sums of class numbers.

Though these methods did not settle Ramanujan's third conjecture, they initiated new and hidden connections with other parts of mathematics. The most spectacular is the 1965 paper of Selberg [180] where he discusses the spectral theory of the Laplace operator and connects it with estimates for $\tau(n)$. In the same paper, he formulates the now celebrated Selberg eigenvalue conjecture of which we shall say more later.

At the end of his paper, Selberg [180] wrote, "It seems most likely that little new may be expected along these lines in the near future, and that a proof of, for instance, the Ramanujan conjecture is more likely to result from future developments in algebraic geometry."

Indeed, if one looks at the Ramanujan conjecture for general Hecke eigenforms, then in 1954, Eichler, Shimura and Igusa solved it for the case $k = 2$ by noting that if we consider

$$\Gamma_0(N) = \left\{ \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}),\ c \equiv 0 \pmod{N} \right\}$$

then $\Gamma_0(N)\backslash\mathfrak{h}$, suitably compactified, has the structure of a Riemann surface and consequently, can be identified as the $\mathbb{C}$-locus of a curve. This curve, denoted $X_0(N)$, is called the modular curve of level N and is defined over the integers. One can therefore consider the zeta function of this curve over the finite field of p elements. The "Riemann hypothesis" for this zeta function turns out to be equivalent to the analogue of the Ramanujan conjecture for weight 2 Hecke eigenforms.

In the early 1960s, it was strongly believed that the Ramanujan conjecture would be proved similarly, not from the "Riemann hypothesis" for curves, but rather from considering higher-dimensional varieties and as a consequence of the Weil conjectures. In fact, Eichler, Shimura, Kuga and Ihara constructed a variety that should have worked but did not quite work since it was not compact. Deligne showed how to compactify their variety and finally, resolved Ramanujan's conjecture. This was the geometric attack on Ramanujan's conjecture, described briefly in the previous chapter.

Yet another analytic method connected to representation theory was outlined by Langlands [108]. We sketch this method below. But first, we have to discuss another aspect of the Ramanujan conjecture, and this is the analogue in the context of Maass wave forms.

2 Maass Forms of Weight Zero

The Ramanujan conjecture and its generalization by Petersson opened the way to a better understanding of what it meant in the bigger picture. Surely, Ramanujan knew that this held in greater generality since in his 1916 paper, he formulated it for other

modular forms as well. This laid the foundation for a conceptual enlightenment. In the 1960s, Jacquet and Langlands [86] reworked Hecke's theory for general automorphic forms on $GL(2)$. Indeed, once the general notion of an automorphic form on $GL(n)$ was precisely formulated, it soon became clear that classical modular forms comprise only half of the theory of automorphic forms on $GL(2)$. There was another side to the story, and this is the theory of Maass wave forms. The Ramanujan–Petersson conjecture for holomorphic Hecke eigenforms could be extended to cover Maass wave forms also, and more generally to automorphic forms on $GL(n)$. This became known as the Ramanujan conjecture for $GL(n)$ and in this generality represents a major unsolved problem in the general theory today. Its solution would have significant consequences in number theory and related areas.

The conceptual insight gained by reformulating the Ramanujan conjecture for $GL(n)$ in this way led to a larger understanding of another celebrated conjecture, namely, the Selberg eigenvalue conjecture, originally formulated in the context of Maass wave forms. In fact, in the adelic formulation, the Ramanujan conjecture "at the infinite prime" becomes the Selberg eigenvalue conjecture. We review this theory briefly below.

Indeed, if we consider modular forms without the holomorphy condition but insist that our function is an eigenfunction of the non-Euclidean Laplacian

$$\Delta = -y^2\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$$

we arrive at the notion of real analytic forms. We may write such a function as a function of the variables x, y, and since $f(z+1) = f(z)$, we have

$$f(x, y) = \sum_n a_n(f, y)e^{2\pi inx}.$$

Suppose that $\Delta f = \lambda f$. This gives us a condition on the coefficients $a_n(f, y)$, namely that they satisfy

$$-y^2\frac{d^2}{dy^2}a_n(f, y) = \left(\lambda - 4\pi^2 n^2 y^2\right)a_n(f, y).$$

This is a Bessel-type differential equation with two solutions. But we only take the one which gives rise to a convergent series above. One can renormalize and show that

$$f(x, y) = a_0(f)y^s + a_0'(f)y^{1-s} + \sum_{n\neq 0} a_n(f)\sqrt{y}K_{ir}\left(2\pi|n|y\right)e^{2\pi inx}$$

where

$$K_{ir}(y) = \frac{1}{2}\int_{-\infty}^{\infty} e^{-y\cosh t - irt}\,dt$$

with $\lambda = 1/4 + r^2$. Maass [117] proved that the series

$$\sum_{n\neq 0} \frac{a_n(f)}{|n|^s}$$

extends to a meromorphic function for all $s \in \mathbb{C}$ analytic everywhere except possibly at $s = 0$ and $s = 1$, and satisfies a functional equation.

The analog of the Ramanujan conjecture in this context is that for any $\epsilon > 0$, $a_n(f) = O(n^\epsilon)$. The Selberg conjecture is that $\lambda \geq 1/4$, or equivalently, r is real and not purely imaginary.

In his 1970 paper, Langlands [108] indicated how the Ramanujan conjecture would follow if a family of L-functions discussed below admit an analytic continuation to a fixed half-plane. At the same time, he interpreted the Selberg conjecture as the Ramanujan conjecture "at infinity" and thus put both conjectures on an equal conceptual footing. (This adelic viewpoint had roots in earlier work of Satake.) We will outline this argument of Langlands below.

At the moment, the best general estimate for $GL(2)$ are due to Kim and Shahidi [97], who showed that $a_n = O(n^{7/64})$, and Kim and Sarnak [95], who showed that $\lambda \geq 0.238$.

3 Upper Bound for Fourier Coefficients and Eigenvalue Estimates

In [108], Langlands also introduced symmetric power L-functions and conjectured that they extend to analytic functions for all values of s. He outlined how this conjecture would lead to a proof of the extended Ramanujan conjecture and the Selberg eigenvalue conjecture. The general theory of symmetric power L-functions has come to play a central role in number theory and representation theory of automorphic forms. We sketch the argument of Langlands below.

We begin with the elementary observation in analytic number theory

$$e^{-1/x} = \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \Gamma(s) x^s \, ds$$

which is easily demonstrated by contour integration and Stirling's formula. Hence,

$$\sum_{n=1}^{\infty} a_n e^{-n/x} = \frac{1}{2\pi i} \int_{(2)} \Gamma(s) f(s) x^s \, ds$$

where

$$f(s) = \sum_{n=1}^{\infty} a_n / n^s.$$

Now suppose that $a_n \geq 0$ and $f(s)$ is absolutely convergent for $\Re(s) \geq 1+\epsilon$. Moving the line of integration to $\Re(s) = 1+\epsilon$ gives

$$\sum_{n=1}^{\infty} a_n e^{-n/x} = O\big(x^{1+\epsilon}\big).$$

Thus, for any individual term in the sum, we have

$$a_n e^{-n/x} = O\big(x^{1+\epsilon}\big).$$

Choosing $x = n$, we deduce that $a_n = O(n^{1+\epsilon})$.

If we let π be an automorphic representation on $GL(2)$ with local parameters, α_p, β_p, then one can give an intuitive description of the symmetric power L-functions attached to π in the following way.

Consider, with r_m denoting the m-th symmetric power of π,

$$L_m(s) := L(s, \pi, r_m) = \prod_p \prod_{j=1}^{m} \left(1 - \frac{\alpha_p^{m-j}\beta_p^j}{p^s}\right)^{-1}$$

where we are ignoring the finitely many Euler factors that need to be modified corresponding to the ramified factors.

Consider the L-function

$$L(s, \pi, r_m \otimes \overline{r}_m) = \prod_{k \leq 2m, k \text{ odd}} L(s, \pi, r_k).$$

The proof of this identity is equivalent to the trigonometric identity

$$1 + \frac{\sin 3\theta}{\sin\theta} + \frac{\sin 5\theta}{\sin\theta} + \cdots + \frac{\sin(2m-1)\theta}{\sin\theta} = \left(\frac{\sin m\theta}{\sin\theta}\right)^2$$

which is easily proved by induction and left as an exercise for the reader.

Thus, the series $L(s, \pi, r_m \otimes \overline{r}_m)$ is a Dirichlet series with non-negative coefficients. If we now suppose that for each $m \geq 1$, $L(s, \pi, r_m)$ is analytic for $\Re(s) \geq 1+\epsilon$, then its pth coefficient (for p prime) is $O(p^{1+\epsilon})$ by the argument given above. But the pth coefficient is easily calculated to be

$$\left|\sum_{j=1}^{m} \alpha_p^{m-j}\beta_p^j\right|^2.$$

Moreover, $|\alpha_p\beta_p| = 1$, so that if the Ramanujan conjecture is false, one of these has absolute value greater than 1. Without any loss of generality, suppose it is α_p. Then, in the above summation, α_p^m dominates the sum, so we deduce

$$|\alpha_p|^{2m} = O\big(p^{1+\epsilon}\big).$$

Taking $2m$th roots, we obtain

$$\alpha_p = O\big(p^{(1+\epsilon)/2m}\big),$$

and letting m tend to infinity, we obtain $\alpha_p = O(1)$, which is the Ramanujan conjecture.

Using the fact that each of the L-functions $L(s, \pi, r_m)$ satisfies a functional equation, one can improve the estimate using a classical result of Chandrasekharan and Narasimhan [32]. This result says that if $a_n \geq 0$ and $f(s) = \sum_{n=1}^{\infty} a_n/n^s$ is convergent in some half-plane, has analytic continuation for all s except for a pole at $s = 1$ of order k and satisfies a functional equation of the form

$$Q^s \Delta(s) f(s) = w Q^{1-s} \Delta(1-s) f(1-s)$$

where $Q > 0$ and

$$\Delta(s) = \prod_i \Gamma(\alpha_i s + \beta_i)$$

then

$$\sum_{n \leq x} a_n = x P_{k-1}(\log x) + O\big(x^{\frac{2A-1}{2A+1}} \log^{k-1} x\big)$$

where $A = \sum_i \alpha_i$. Taking differences, we deduce that

$$a_n = O\big(n^{\frac{2A-1}{2A+1}} \log^{k-1} n\big).$$

In [129], this result is stated with a typo on page 525. (On lines 3 and 7 of [129], $(2A-1)(2A+1)$ should be $(2A-1)/(2A+1)$ in both instances.) Hardy [67] was aware of this argument, at least for $m = 2$, as is indicated in the tenth of his twelve lectures on Ramanujan's work.

A similar reasoning can be applied to obtain bounds in the Selberg eigenvalue conjecture. If π corresponds to a Maass form with eigenvalue λ, then the Gamma factors in the functional equation of $L(s, \pi, r_m)$ will have the following shape:

$$\Gamma(s, \pi, r_m) = \prod_{j=0}^{m} \Gamma\left(\frac{s - \lambda_j}{2}\right), \quad \lambda_j = i(m-2j)r, \quad \lambda = \frac{1}{4} + r^2.$$

Using the conjectured analyticity of $L_m(s)$ for all m and this information on the Γ-factor for $L_m(s)$ now gives the Selberg eigenvalue conjecture.

4 Eisenstein Series

It is interesting to note that R.A. Rankin was a doctoral student of G.H. Hardy and had taken up Ramanujan's conjecture for his doctoral thesis in 1933. The thesis was

a fundamental advance since it set the stage to what was to become the celebrated Rankin–Selberg method. This method involved an unfolding trick that generalized to the $GL(n)$ setting as was noted by Jacquet, Piatetski-Shapiro and Shalika [87].

We present this method in its simplest setting. The earliest example of a Maass form is given by the Eisenstein series

$$E(z,s) = \pi^{-s}\Gamma(s)\frac{1}{2}\sum_{(m,n)\neq(0,0)}\frac{y^s}{|mz+n|^{2s}}.$$

This series converges for $\Re(s) > 1$, and we clearly have

$$E(\gamma z, s) = E(z,s)$$

for all $\gamma \in SL_2(\mathbb{Z})$. In addition, it is easily verified that

$$\Delta E(z,s) = s(1-s)E(z,s)$$

so that $E(z,s)$ is a weight zero Maass form with eigenvalue $s(1-s)$. Since $E(z,s)$ is periodic with period 1, we can derive its Fourier series:

$$E(z,s) = \sum_{r=-\infty}^{\infty} a_r(y,s)e^{2\pi irx}$$

and

$$a_r(y,s) = \int_0^1 E(x+iy,s)e^{-2\pi irx}\,dx.$$

We do the obvious. We insert the series expansion for $E(z,s)$ into the integral and apply Fubini's theorem. First, the contribution to $E(z,s)$ from $m=0$ is

$$\pi^{-s}\Gamma(s)y^s\zeta(2s).$$

This is part of $a_0(y,s)$ but not all of a_0 as we shall see below. Now suppose $m \neq 0$. Since (m,n) and $(-m,-n)$ give the same summand in $E(z,s)$, we may suppose that $m > 0$. Thus,

$$a_r(y,s) = \pi^{-s}\Gamma(s)y^s\sum_{m=1}^{\infty}\sum_{n=-\infty}^{\infty}\int_0^1\left[(mx+n)^2+m^2y^2\right]^{-s}e^{-2\pi irx}\,dx.$$

If we put $n = qm + d$ with $0 \le d < m$, the sum becomes

$$\sum_{m=1}^{\infty}\sum_{d \bmod m}\int_{-\infty}^{\infty}\left[(mx+d)^2+m^2y^2\right]^{-s}e^{-2\pi irx}\,dx.$$

We change the variable $x = u - d/m$ to get

$$\sum_{m=1}^{\infty} m^{-2s} \int_{-\infty}^{\infty} \left(u^2+y^2\right)^{-s} e^{-2\pi i r u} \left(\sum_{d \bmod m} e^{2\pi i d r/m} \right) du.$$

The innermost sum is zero unless $m|r$ in which case it is m. Thus, the sum becomes

$$\sum_{m|r} m^{1-2s} \int_{-\infty}^{\infty} \left(u^2+y^2\right)^{-s} e^{-2\pi i r u}\, du.$$

If $r = 0$, we get

$$\pi^{-s}\Gamma(s) y^s \zeta(2s-1) \int_{-\infty}^{\infty} \left(u^2+y^2\right)^{-s} du$$

which is equal to

$$\pi^{-s}\sqrt{\pi}\,\Gamma(s-1/2) y^{1-s} \zeta(2s-1).$$

Thus, the constant term (on applying the functional equation for $\zeta(s)$) is equal to

$$a_0(y,s) = \pi^{-s}\Gamma(s)\zeta(2s) y^s + \pi^{s-1}\Gamma(1-s)\zeta(2-2s) y^{1-s}.$$

If $r \neq 0$, then we get

$$a_r(y,s) = 2|r|^{s-1/2} \sigma_{1-2s}\left(|r|\right) \sqrt{y} K_{s-1/2}\left(2\pi |r| y\right)$$

where

$$\sigma_{1-2s}(r) = \sum_{m|r} m^{1-2s}.$$

One can show that $a_r(y,s) = a_r(y,1-s)$ and $r^s \sigma_{-2s}(r) = r^{-s}\sigma_{2s}(r)$, from which the functional equation is easily deduced.

5 Eisenstein Series and Non-vanishing of $\zeta(s)$ on $\Re(s)=1$

We want to indicate a proof of the non-vanishing of $\zeta(s)$ on $\Re(s)=1$ which uses the theory of Eisenstein series and as a consequence does not use the Euler product of $\zeta(s)$ as most conventional proofs do. The idea was used by Jacquet and Shalika [89] in their general result about the non-vanishing on $\Re(s)=1$ of automorphic L-functions associated with GL_n.

Recall that

$$E(z,s) = \pi^{-s}\Gamma(s)\frac{1}{2} \sum_{(m,n)\neq(0,0)} \frac{y^s}{|mz+n|^{2s}}.$$

Notice that we may also write this as

$$E(z,s) = \pi^{-s}\Gamma(s)\frac{1}{2}\zeta(2s)\sum_{\gamma\in\Gamma_\infty\backslash\Gamma}\Im(\gamma z)^s$$

where Γ_∞ is the stabilizer of the cusp at infinity.

We have already indicated that

$$\begin{aligned}E(z,s) = {}& \pi^{-s}\Gamma(s)\zeta(2s)y^s + \pi^{s-1}\Gamma(1-s)\zeta(2-2s)y^{1-s}\\ &+ \sum_{r\neq 0}|r|^{s-1/2}\sigma_{1-2s}\big(|r|\big)\sqrt{y}K_{s-1/2}\big(2\pi|r|y\big)e^{2\pi irx}\end{aligned}$$

where $\sigma_v(n) = \sum_{d|n} d^v$ and

$$K_s(y) = \frac{1}{2}\int_0^\infty e^{-y(t+t^{-1})/2}t^s\,\frac{dt}{t}.$$

One can prove directly that $K_s(y) = K_{-s}(y)$ and $r^s\sigma_{-2s}(r) = r^s\sigma_{2s}(r)$, which allows us to deduce the functional equation of $E(z,s)$ from its Fourier expansion.

This result lies at the heart of the Langlands–Shahidi method of analytic continuation of Eisenstein series. It is also at the core of the Rankin–Selberg method of analytic continuation discussed in the next section.

Now suppose that $\zeta(1+it_0) = 0$ for some t_0 real. Then, $\zeta(1-it_0) = 0$ also. We put $s = (1+t_0)/2$ in $E(z,s)$. Then, the constant term vanishes, and we get a Maass cusp form:

$$E\big(z,(1+it_0)/2\big) = 4\sqrt{y}\sum_{r=1}^\infty r^{it_0/2}\sigma_{-it_0}(r)\cos(2\pi rx)\int_0^\infty e^{-\pi ry(t+t^{-1})}\,\frac{dt}{t^{1-it_0/2}}.$$

Using standard estimates for the integral, one can show that the sum is $O(e^{-cy})$ for some $c > 0$. As the constant term of $E(z,(1+it_0)/2)$ is zero, we have a genuine Maass cusp form on our hands.

In particular,

$$\int_0^1 E\big(x+iy,(1+it_0)/2\big)\,dx = 0.$$

Multiplying this equation by y^{s-2} and integrating from 0 to ∞, we get

$$\int_0^\infty\int_0^1 E\big(x+iy,(1+it_0)/2\big)y^{s-2}\,dx\,dy = 0.$$

Now we use the basic idea that

$$\bigcup_{\gamma\in\Gamma_\infty\backslash\Gamma}\gamma(\Gamma\backslash H) = [0,1]\times[0,\infty],$$

usually referred to as the "unfolding" of the domain of integration. Thus,

$$\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \int_{\gamma(\Gamma \backslash H)} E\big(z, (1 + it_0)/2\big) \Im(z)^s \, \frac{dx\, dy}{y^2} = 0.$$

As $E(\gamma z, s) = E(z, s)$, we may change variables and get:

$$\begin{aligned} 0 &= \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \int_{\Gamma \backslash H} E\big(z, (1 + it_0)/2\big) \Im(\gamma z)^s \, \frac{dx\, dy}{y^2} \\ &= \int_{\Gamma \backslash H} E\big(z, (1 + it_0)/2\big) E(z, s) \, \frac{dx\, dy}{y^2} \end{aligned}$$

valid for all $s \in \mathbb{C}$.

From the definition of $E(z, s)$ (or its Fourier expansion) we see that

$$E(z, \overline{s}) = \overline{E(z, s)}.$$

Therefore, putting $s = (1 - it_0)/2$, we get from the penultimate equation:

$$0 = \int_{\Gamma \backslash H} \big|E\big(z, (1 + it_0)/2\big)\big|^2 \, \frac{dx\, dy}{y^2}.$$

Thus, the integrand is identically zero. That is, we have proved that $\zeta(1 + it_0) = 0$ implies that

$$E\big(z, (1 + it_0)/2\big) \equiv 0.$$

We now show that this is a contradiction. We do this by showing that some Fourier coefficient of $E(z, (1 + it_0)/2)$ is non-zero. That is, we need to check

$$\int_0^\infty e^{-\pi r y (u + u^{-1})} \, \frac{du}{u^{1 + it_0}} \neq 0.$$

If we set $u = e^\theta$, we have to show that

$$\int_{-\infty}^\infty e^{-\pi r y (e^\theta + e^{-\theta}) - it_0 \theta} \, d\theta \neq 0.$$

In other words, it suffices to show that

$$\int_0^\infty e^{-\pi r y (e^\theta + e^{-\theta})} \cos t_0 \theta \, d\theta \neq 0.$$

This integral is of the form

$$\int_0^\infty e^{-y(a^\theta + a^{-\theta})} \cos \theta \, d\theta, \quad a > 1.$$

We would like to determine its behaviour as y tends to infinity. To do this, we can apply Laplace's saddle point method: if f has two continuous derivatives, with $f(0) = f'(0) = 0$ and $f''(0) > 0$, and f is increasing in $[0, A]$, then

$$I(x) := \int_0^A e^{-xf(t)}\,dt \sim \sqrt{\frac{\pi}{2xf''(0)}}$$

as x tends to infinity and provided that $I(x_0)$ exists for some x_0. A slightly generalized version of this says that if g is continuous on $[0, A]$, then

$$\int_0^A g(t)e^{-xf(t)}\,dt \sim g(0)\sqrt{\frac{\pi}{2xf''(0)}}.$$

Now choose $f(t) = a^t + a^{-t} - 2$, $g(t) = \cos t$ so that

$$e^{-2x}\int_0^\infty e^{-x(a^\theta + a^{-\theta} - 2)}\cos\theta\,d\theta \sim e^{-2x}2\log a\sqrt{\frac{\pi}{x}}$$

from which we see that $E(z, (1 + it_0)/2) \not\equiv 0$, as required. This gives the desired contradiction. It is possible to deduce the non-vanishing of the above integrals directly without appealing to Laplace's saddle point method. With some work, it may also be possible to derive a zero-free region for $\zeta(s)$.

6 The Rankin–Selberg L-Function

The unfolding technique employed in the previous section has wider ramifications. It can be used to establish the analytic continuation and functional equation for a large class of L-functions which fall under the umbrella of Rankin–Selberg theory.

Let $F : H \to \mathbb{C}$ be a Γ-invariant function which is of rapid decay (that is, $F(x + iy) = O(y^N)$ for all $N \geq 1$.) Let

$$C(F, y) = \int_0^1 F(x + iy)\,dx, \quad y > 0$$

be the constant term of the Fourier expansion. Let

$$L(F, s) = \int_0^\infty C(F, y)y^s\,\frac{dy}{y^2}$$

be the Mellin transform of $C(F, y)$.

Theorem 6.1 *Let $L^*(F, s) = \pi^{-s}\Gamma(s)\zeta(2s)L(F, s)$. Then, $L(F, s)$ has analytic continuation to the whole complex plane, regular everywhere except for a simple pole at $s = 1$ with residue equal to*

$$\frac{3}{\pi}\int_{\Gamma\backslash H} F(z)\,dz.$$

The function $L^*(F,s)$ is regular for all $s \neq 0,1$ and satisfies the functional equation

$$L^*(F,s) = L^*(F,1-s).$$

Proof The key idea is to use the decomposition described earlier. We have

$$L(F,s) = \int_0^\infty \int_0^1 F(x+iy) y^{s-2}\, dx\, dy.$$

Decomposing the domain of integration as in the "unfolding" technique, this becomes

$$= \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \int_{\gamma(\Gamma \backslash H)} F(z) y^s \frac{dx\, dy}{y^2}.$$

This can be rewritten as

$$\sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \int_{\Gamma \backslash H} F(\gamma z) \big(\Im(\gamma z)\big)^s \frac{dx\, dy}{y^2} = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} \int_{\Gamma \backslash H} F(z) \big(\Im(\gamma z)\big)^s \frac{dx\, dy}{y^2}$$

because F is Γ-invariant. Moving the summation inside the integral shows that this is equal to

$$\int_{\Gamma \backslash H} F(z) E(z,s) \frac{dx\, dy}{y^2}.$$

As $E(z,s)$ has analytic continuation and functional equation, we get the same for $L(F,s)$. □

We now give a few examples on how to apply this theorem.

In the special case that f is a cusp form of weight k, we may apply the above result to $F(z) = y^k |f(z)|^2$, which is easily checked to be Γ-invariant.

A straightforward computation shows that the constant term is

$$y^k \sum_{n=1}^\infty |a_n|^2 e^{-4\pi n y}.$$

The Mellin transform of the constant term is

$$\int_0^\infty y^{k+s} \sum_{n=1}^\infty |a_n|^2 e^{-4\pi n y} \frac{dy}{y^2} = (4\pi)^{-s-k+1} \Gamma(s+k-1) \sum_{n=1}^\infty \frac{|a_n|^2}{n^{s+k-1}}.$$

This proves the following:

Theorem 6.2 *Let f be a cusp form of weight k for $SL_2(\mathbb{Z})$. If*

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$$

is its Fourier expansion at infinity, then the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{|a_n|^2}{n^s}$$

has a meromorphic continuation to the whole complex plane. In fact, if

$$\psi(s) = \pi^{-2s-k+1} 2^{-2s} \Gamma(s)\Gamma(s+k-1)\zeta(2s) \sum_{n=1}^{\infty} \frac{|a_n|^2}{n^s}$$

then $\psi(s)$ extends to a function which is regular for all $s \in \mathbb{C}$ except at $s = 1$ where it has a simple pole and residue equal to

$$\frac{3}{\pi} \int_{\Gamma \backslash H} y^k \left| f(z) \right|^2 \frac{dx\, dy}{y^2} = \frac{3}{\pi}(f, f).$$

Moreover, $\psi(s)$ satisfies the functional equation $\psi(s) = \psi(1-s)$.

If we apply the theorem of Chandrasekharan and Narasimhan [32] mentioned in an earlier section, we deduce that

$$\sum_{n \le x} |a_n|^2 = \frac{3}{\pi}(f, f) x^k + O\left(x^{k-2/5}\right)$$

because twice the sum of the coefficients in the Gamma factors (or equivalently the degree in the sense of Selberg) is equal to 4. By taking a single summand in the sum on the left, we deduce that $a_n = O(n^{k/2-1/5})$. The same technique applied to Maass forms gives us $a_n = O(n^{3/10})$.

If we take f and g to be cusp forms (or even with one of them a cusp form), we consider

$$y^k f(z) \overline{g(z)}$$

which is Γ-invariant. If

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi i n z}$$

and

$$g(z) = \sum_{n=1}^{\infty} b_n e^{2\pi i n z}$$

are the respective Fourier expansions at infinity, then the constant term is easily computed to be equal to

$$y^k \sum_{n=1}^{\infty} a_n \overline{b_n} e^{-4\pi n y}.$$

One could also take forms of different weights k_1 and k_2 and consider

$$y^{(k_1+k_2)/2} f(z)\overline{g(z)}.$$

In the end, applying Theorem 6.1, we deduce that

$$L_{f,g}(s) := \sum_{n=1}^{\infty} \frac{a_n \overline{b_n}}{n^s}$$

admits an analytic continuation to the entire complex plane. If f and g are normalized Hecke eigenforms, then the series has an Euler product. This is essentially a consequence of Ramanujan's identity

$$\sum_{n=0}^{\infty} \left(\frac{z^{n+1}-1}{z-1}\right)\left(\frac{w^{n+1}-1}{w-1}\right) q^n = \frac{1-zwq^2}{(1-q)(1-zq)(1-wq)(1-zwq)}.$$

A suitably normalized version of this series $L_{f,g}(s)$ (with appropriate Γ-factors, $\zeta(2s)$ and so forth) extends to a function which is regular everywhere except possibly at $s = 1$ where it may have a simple pole with residue equal to

$$\frac{3}{\pi}(f, g).$$

Thus, if f and g are orthogonal to each other, then the normalized series extends to an entire function.

Kronecker's limit formula states that

$$\lim_{s\to 1}\left[E(z, s) - \frac{1}{s-1}\right] = \log\left(e^{\gamma}/4\pi\right) - 2\log\left(\sqrt{y}\left|\eta(z)\right|^2\right)$$

where $\eta(z) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n)$ with $q = e^{2\pi i z}$. If f and g are Hecke eigenforms with π_f, π_g being the associated automorphic representations, the Kronecker limit formula allows us to write down an exact formula for the special value $L(1, \pi_f \otimes \pi_g)$.

7 Poincaré Series for $SL_2(\mathbb{Z})$

The Poincaré series for $SL_2(\mathbb{Z})$ are defined by

$$G_r(z) = \frac{1}{2} \sum_{(c,d)=1} (cz+d)^{-k} e^{2\pi i r \frac{az+b}{cz+d}}$$

where the summation is over all integers c, d with $(c, d) = 1$, and a, b are any integers such that $ad - bc = 1$. Observe that if $r = 0$, this reduces to the classical Eisenstein series $E_k(z)$ (up to a constant). Thus, the Poincaré series are to be viewed as generalizations of Eisenstein series. It is easy to see that the inner summand does not depend on the choice of a solution. Indeed, by the Euclidean algorithm, any other solution for (a, b) has the form $(a + tc, b + td)$, and

$$\frac{(a+tc)z+(b+td)}{cz+d} = \frac{az+b}{cz+d} + t, \quad t \in \mathbb{Z}$$

so that

$$e^{2\pi i r(\frac{az+b}{cz+d}+t)} = e^{2\pi i r(\frac{az+b}{cz+d})}.$$

We can rewrite the series in a more invariant form by setting

$$j(\gamma, z) = cz + d, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

and then

$$G_r(z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, z)^{-k} e^{2\pi i r(\gamma z)}.$$

The important thing to note is that $G_r(z)$ is a modular form of weight k. To see this, let $\delta \in \Gamma = SL_2(\mathbb{Z})$. Then,

$$G_r(\delta z) = \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma, \delta z)^{-k} e^{2\pi r(\gamma \delta z)}.$$

Now, we have the so-called cocycle relation

$$j(\gamma\delta, z) = j(\gamma, \delta z) j(\delta, z)$$

as is easily verified, so that

$$j(\gamma, \delta z) = \frac{j(\gamma\delta, z)}{j(\delta, z)}$$

and

$$\begin{aligned} G_r(\delta z) &= j(\delta, z)^k \sum_{\gamma \in \Gamma_\infty \backslash \Gamma} j(\gamma\delta, z)^{-k} e^{2\pi i r(\gamma\delta z)} \\ &= j(\delta, z)^k G_r(z). \end{aligned}$$

Holomorphy is easy to verify using standard tests of complex analysis. In addition, $G_r(i\infty) = 0$ if $r \geq 1$. We conclude that for every $r \geq 1$, $G_r(z)$ is a cusp form of weight k. Thus, Poincaré series give explicit constructions of cusp forms.

For a detailed treatment of this theory, we refer the reader to Rankin's book [171] (especially Chap. 5).

Now let f be any cusp form of weight k. We would like to compute the inner product (f, G_r).

First observe that $e^{2\pi iz} = e^{2\pi ix} \cdot e^{-2\pi y}$ so that

$$\overline{e^{2\pi iz}} = e^{-2\pi ix} \cdot e^{-2\pi y} = e^{-2\pi i(\overline{z})}.$$

Thus,

$$\begin{aligned}
(f, G_r) &= \int_{\Gamma\backslash H} \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} f(z)e^{-2\pi ir\overline{\gamma z}}(\overline{cz+d})^{-k}y^k \frac{dx\,dy}{y^2} \\
&= \int_{\Gamma\backslash H} \sum_{\Gamma_\infty\backslash\Gamma} (cz+d)^k f(z)e^{-2\pi ir(\overline{\gamma z})} \frac{y^k}{|cz+d|^{2k}} \frac{dx\,dy}{y^2} \\
&= \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\Gamma\backslash H} f(\gamma z)\Im(\gamma z)^k e^{-2\pi ir(\overline{\gamma z})} \frac{dx\,dy}{y^2} \\
&= \sum_{\gamma\in\Gamma_\infty\backslash\Gamma} \int_{\gamma(\Gamma\backslash H)} f(z)\Im(z)^k e^{-2\pi ir\bar{z}} \frac{dx\,dy}{y^2} \\
&= \int_0^\infty \int_0^1 f(x+iy)y^k e^{-2\pi irx}e^{-2\pi ry} \frac{dx\,dy}{y^2}.
\end{aligned}$$

Now from the Fourier expansion of $f(z)$ we have

$$f(z) = \sum_{n=1}^{\infty} a_n e^{2\pi inx} \cdot e^{-2\pi ny}.$$

We see that the x-integral picks up the rth Fourier coefficient. Thus,

$$(f, G_r) = a_r \int_0^\infty e^{-4\pi ry} y^{k-2}\, dy.$$

By setting $4\pi ry = t$ in the integrand and simplifying, we deduce the following:

Theorem 7.1 *Let f be any cusp form of weight k for Γ. Then,*

$$(f, G_r) = \frac{\Gamma(k-1)a_r}{(4\pi r)^{k-1}}.$$

An important corollary is:

Corollary 7.2 *Every cusp form is a finite linear combination of Poincaré series $G_r(z)$.*

Proof The set of Poincaré series spans a closed subspace in the space of cusp forms. If f is a cusp form not in this space, all of its Fourier coefficients must vanish by the previous theorem. Thus, the orthogonal complement is zero. □

As an example, consider the case $k = 12$. Each of the $G_r(z)$ is a cusp form of weight 12. But any cusp form of weight 12 must be a constant multiple of Δ, Ramanujan's cusp form. Thus,

$$G_r(z) = c_r \Delta(z).$$

What is c_r? By the above theorem,

$$(\Delta, G_r) = \frac{\Gamma(11)\tau(r)}{(4\pi r)^{11}} = c_r(\Delta, \Delta).$$

Hence,

$$\tau(r) = \frac{(4\pi r)^{11}}{10!} \int_{\Gamma\backslash H} y^{12} \Delta(z) \overline{G_r(z)} \frac{dx\, dy}{y^2}$$

which is an interesting formula for the Ramanujan τ-function as an inner product.

8 Fourier Coefficients and Kloosterman Sums

Emanating largely from the work of Petersson [154] in the 1930s and Selberg [179], an explicit formula can be given for the Fourier coefficients of $G_r(z)$. This striking formula involves the Kloosterman sums, and their appearance has opened new connections to the Selberg eigenvalue conjecture as well as applications to classical questions of analytic number theory. We now derive this remarkable formula.

We begin by writing

$$G_r(z) = \sum_{n=1}^{\infty} g_{rn} e^{2\pi i n z}.$$

Then,

$$g_{rn} = \int_0^1 G_r(x) e^{-2\pi i n x}.$$

More precisely, for reasons of convergence, we should consider

$$\int_{i\delta}^{1+i\delta} G_r(z) e^{-2\pi i n z}$$

with $\delta > 0$, but we leave this technical modification to the reader. We have

$$g_{rn} = \frac{1}{2} \sum_{(c,d)=1} \int_0^1 (cx + d)^{-k} e^{2\pi i r(\frac{ax+b}{cx+d}) - 2\pi i n x}\, dx.$$

Put $cx + d = t$. The argument in the exponential becomes

$$\frac{r}{t}\left(\frac{a}{c}(t-d)+b\right)-\frac{n}{c}(t-d)=\frac{nd+ar}{c}-\frac{nt}{c}-\frac{r}{tc}$$

since $ad - bc = 1$. Thus,

$$g_{rn}=\frac{1}{2}\sum_{c\neq 0}\frac{1}{c}\sum_{\substack{d(\mathrm{mod}\ c)\\ ad\equiv 1(\mathrm{mod}\ c)}} e^{\frac{2\pi i}{c}(nd+ar)}\int_{-\infty}^{\infty} t^{-k}e^{-\frac{2\pi i}{c}(\frac{r}{t}+nt)}\,dt$$

because

$$\sum_{(c,d)=1}\int_0^1 t^{-k}e^{\frac{2\pi i}{c}(nd+ar-nt)}e^{-\frac{2\pi ir}{t}}\,dt$$

depends only on $d \pmod c$. Writing d as $d_0+(m+1)c$ with varying m, we transform the integral from 0 to 1 into an integral from $-\infty$ to ∞. This integral turns out to be a Bessel function:

$$\int_{-\infty+ci}^{\infty+ci} t^{-k}\exp\left(-\frac{2\pi i}{c}\left(\frac{r}{t}+nt\right)\right)dt = 2\pi(n/r)^{(k-1)/2}J_{k-1}(4\pi\sqrt{rn}/c)$$

where

$$J_k(z)=\frac{1}{2\pi i}\int_C t^{-k-1}e^{\frac{z}{2}(t-1/t)}\,dt$$

where C is the unit circle. The sum

$$S(r,n,c):=\sum_{\substack{d(\mathrm{mod}\ c)\\ ad\equiv 1(\mathrm{mod}\ c)}} e^{\frac{2\pi i}{c}(nd+ar)}$$

is called a Kloosterman sum. Using this notation, we obtain the beautiful formula due to Petersson:

$$g_{rn}=(n/r)^{(k-1)/2}\left\{\delta_{rn}+\pi\sum_{c=1}^{\infty}\frac{S(r,n,c)}{c}J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right)\right\}$$

where δ_{rn} denotes the Kronecker delta function.

We have already noted that the Poincaré series span the space of cusp forms. Thus, to prove the Ramanujan conjecture, it suffices to show that

$$g_{rn}=O\left(n^{\frac{k-1}{2}+\epsilon}\right)$$

for every r. This is tantamount to showing that the expression in parentheses in the above sum is $O(n^{\epsilon})$.

Selberg, using this expression and Weil's estimate for Kloosterman sums:

$$\left|S(r,n,c)\right|\leq d(c)c^{1/2}(r,n,c)^{1/2}$$

as well as the bound

$$J_{k-1}(x) \le A \min\bigl(x^{k-1}, x^{-1/2}\bigr)$$

obtained that

$$g_{rn} = O\bigl(n^{k/2-1/4+\epsilon}\bigr).$$

Note that this is better than what we obtained earlier by the Rankin–Selberg method. Since the estimate was obtained crudely, Selberg felt that there must be cancellation among the Kloosterman sums. This led him to formulate the following conjecture (which was also arrived at independently by Linnik).

Conjecture (Selberg–Linnik)

$$G(x) := \sum_{c \le x} \frac{S(r,n,c)}{c} = O\bigl(x^{\epsilon}\bigr)$$

for $x \ge \gcd(r,n)^{1/2+\epsilon}$ *for any* $\epsilon > 0$.

In his 1965 paper, Selberg stated that this would lead to a proof of the Ramanujan conjecture (for Maass forms as well) but did not indicate a proof. We will indicate below how such a proof can be obtained for the full modular group. The argument is adapted from a paper of R. Murty [129].

Let us first observe that Weil's estimate for Kloosterman sums leads to the estimate

$$G(x) = O\bigl(x^{1/2}\log x\bigr)$$

for $(r,n) = 1$. Kuznetsov [102] proved that $G(x) = O(x^{1/6+\epsilon})$, but the O-constant depends on r, n, and so it is not applicable to the estimation of the Fourier coefficients. Let

$$H(x) := \sum_{c \le x} S(r,n,c).$$

By partial summation, the Selberg–Linnik conjecture is equivalent to

$$H(x) = O\bigl(x^{1+\epsilon}\bigr).$$

We begin by considering

$$\sum_{c>\sqrt{n}} \frac{S(r,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right)$$

$$= \sum_{c>\sqrt{n}} G(c)\left\{J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c+1}\right) - J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right)\right\}$$

by partial summation. By the mean value theorem, the expression in parentheses is

$$\frac{4\pi\sqrt{rn}}{c(c+1)}J'_{k-1}(\xi_c)$$

for some $\xi_c \in (4\pi\sqrt{rn}/(c+1), 4\pi\sqrt{rn}/c)$. Using the estimate

$$J'_{k-1}(x) \ll x^{-1/2}$$

we get

$$\sum_{c>\sqrt{rn}} \frac{S(r,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right) \ll n^{1/4} \sum_{c>\sqrt{n}} \frac{|G(c)|}{c^{3/2}} \ll n^{\epsilon}$$

by the Selberg–Linnik conjecture. Thus, we need only consider

$$\sum_{c\le\sqrt{n}} \frac{S(r,n,c)}{c} J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right).$$

To estimate this, we apply an inductive argument. As there are no cusp forms of weight 10, we have

$$\sum_{c\le\sqrt{n}} \frac{S(r,n,c)}{c} J_9\left(\frac{4\pi\sqrt{rn}}{c}\right) = O\left(n^{\epsilon}\right).$$

So, if for example, we were trying to establish the conjecture for $k = 12$, then it suffices to estimate for $k = 10$ the quantity

$$\sum_{c\le\sqrt{n}} \frac{S(r,n,c)}{c} \left\{ J_{k+1}\left(\frac{4\pi\sqrt{rn}}{c}\right) + J_{k-1}\left(\frac{4\pi\sqrt{rn}}{c}\right)\right\}.$$

By the familiar identity

$$\frac{2kJ_k(x)}{x} = J_{k+1}(x) + J_{k-1}(x)$$

it suffices to estimate

$$\frac{1}{\sqrt{n}} \sum_{c\le\sqrt{n}} S(r,n,c) J_k\left(\frac{4\pi\sqrt{rn}}{c}\right).$$

Again, by partial summation, we may write this as

$$\frac{1}{\sqrt{n}} \sum_{c\le\sqrt{n}} H(c) \left\{ J_k\left(\frac{4\pi\sqrt{rn}}{c+1}\right) - J_k\left(\frac{4\pi\sqrt{rn}}{c}\right)\right\}.$$

Again, the expression in the brackets is

$$\frac{4\pi\sqrt{rn}}{c(c+1)}J_k'(\xi_c).$$

Using the estimate

$$J_k'(x) \ll x^{-1/2}$$

as before and the fact that $H(c) = O(c^{1+\epsilon})$, we deduce a final estimate of $O(n^\epsilon)$ as desired. This completes the proof of the fact that the Selberg–Linnik conjecture implies the Ramanujan conjecture (for the full modular group). A similar argument can be applied to higher levels. However, the non-existence of cusp forms of small weight is not guaranteed. In this case, we exploit the fact that we know the Ramanujan conjecture in the weight two case (a result due to Eichler and Shimura).

9 The Kloosterman–Selberg Zeta Function

In order to gain more insight into the Selberg–Linnik conjecture, we will consider (with Selberg [180]) the series

$$Z(r,n,s) = \sum_{c\neq 0}\frac{S(r,n,c)}{|c|^{2s}}.$$

To study this series, Selberg [178] considers the cognate Poincaré series

$$U_n(z,s) = \sum_{\Gamma_\infty\backslash\Gamma}\Im(\gamma z)^s e^{2\pi i n\gamma z}.$$

Clearly, U_n is Γ-invariant. Moreover, if

$$\Delta = -y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)$$

then

$$\Delta U_n(z,s) = s(1-s)U_n(z,s) + 4\pi n U_n(z,s+1)s.$$

One can show that the Fourier expansion of $U_n(z,s)$ contains $Z(r,n,s)$ in it. This allows us to relate the eigenvalues of Δ with the abscissa of convergence of $Z(r,n,s)$. More precisely,

$$U_n(z,s) = \sum_{m=-\infty}^{\infty} B_n(m,y,s)e^{2\pi imx}$$

where

$$B_n(m, y, s) = \delta_{nm} y^s e^{-2\pi ny} + \frac{1}{2} \sum_{c \neq 0} \frac{S(m, n, c)}{|c|^{2s}} y^{1-s} \times \int_{-\infty}^{\infty} \exp\left(-2\pi imyv - \frac{2\pi n}{c^2 y(1 - iv)}\right) \frac{dv}{(1 + v^2)^s}.$$

It then turns out that

$$(2\pi\sqrt{nm})^{2s-1} \sum_{c=1}^{\infty} \frac{S(n, m, c)}{|c|^{2s}} = \frac{\sin \pi s}{2} \sum_{j=1}^{\infty} \frac{a_j(n)\overline{a_j(m)}}{\cosh \pi r_j} \Gamma\left(s - \frac{1}{2} + ir_j\right) \Gamma\left(s - \frac{1}{2} - ir_j\right) - \frac{\delta_{nm}}{2\pi} \frac{\Gamma(s)}{\Gamma(1 - s)} + \frac{1}{\pi} \int_{-\infty}^{\infty} (n/m)^{ir} \sigma_{2ir}(m) \sigma_{-2ir}(m) \frac{h(r, s)}{|\zeta(1 + 2ir)|^2} dr$$

where

$$h(r, s) = \frac{\sin \pi s}{2} \Gamma\left(s - \frac{1}{2} + ir\right) \Gamma\left(s - \frac{1}{2} - ir\right)$$

and the $a_j(n)$s are the Fourier coefficients of the Maass form corresponding to the eigenvalue $\lambda_j = 1/4 + r_j^2$. This remarkable formula establishes a striking relationship between the eigenvalues λ_j and the Kloosterman–Selberg zeta function.

10 Rankin–Selberg L-Functions for GL_n

Let $GL_2^+(\mathbb{R})$ be the group of 2×2 matrices with real entries and positive determinant. The upper half-plane can be realized as a symmetric space of $GL_2^+(\mathbb{R})$. More precisely, if Z denotes the subgroup of scalar matrices and $SO(2)$ denotes the special orthogonal group, then the upper half-plane can be realized as the coset space $GL_2^+(\mathbb{R})/ZSO(2)$. Modular forms then can be viewed as functions on $GL_2^+(\mathbb{R})$ that transform appropriately under certain subgroups. Once we view the classical theory from this perspective, it affords a vast generalization. We refer the reader to [36] for details.

For the moment, it will suffice to say that classical Hecke eigenforms correspond to certain representations of $GL_2(\mathbb{A}_{\mathbb{Q}})$ where $\mathbb{A}_{\mathbb{Q}}$ denotes the adele ring of the rational number field. More generally, one can consider cuspidal automorphic representations of $GL_n(\mathbb{A}_Q)$. Such representations π can be decomposed as a restricted tensor product $\pi_\infty \otimes (\otimes' \pi_p)$, and one can associate parameters $\mu_{1,\infty}, \ldots, \mu_{n,\infty}$ to π_∞, and $\alpha_{1,p}, \ldots, \alpha_{n,p}$ called Satake parameters (for all but finitely many primes p). The Ramanujan conjecture for GL_n is the prediction that $|\alpha_{i,p}| = 1$ for all primes p and

all $1 \leq i \leq n$. The Selberg eigenvalue conjecture translates into the assertion that $\Re(\mu_{j,\infty}) = 0$ for $1 \leq j \leq n$.

Many of the methods we discussed in the earlier sections can be developed to deal with the higher rank case. We refer the reader to [36] for further details.

With reference to the Rankin–Selberg method, the general theory for GL_n was initiated and developed by Jacquet, Piatetski-Shapiro and Shalika [87], Shahidi [187] and finally completed by Moeglin and Waldspurger [125]. If π_1 and π_2 are cuspidal automorphic representations of GL_m and GL_n of the adele ring over the rationals (say), then the Rankin–Selberg L-function is defined by the Euler product

$$L(s, \pi_1 \otimes \pi_2) = \prod_p L(s, \pi_{1,p} \otimes \pi_{2,p})$$

where for all but finitely many primes p, the Euler factors are given by the formula

$$L(s, \pi_{1,p} \otimes \pi_{2,p}) = \prod_{i,j} \left(1 - \frac{\alpha_{i,p}^{(1)} \alpha_{j,p}^{(2)}}{p^s}\right)^{-1}$$

and

$$L(s, \pi_{r,p}) = \prod_i \left(1 - \frac{\alpha_{i,p}^{(r)}}{p^s}\right)^{-1}$$

for $r = 1, 2$. It is possible to define the Euler factors at all the places so that the final product converges for $\Re(s) > 1$. The completed L-function turns out to be entire unless

$$\pi_2 \simeq \pi_1^v \otimes |\det|^{it}$$

for some real number t, in which case the function is regular everywhere except at $s = 1 - it$ where it has a simple pole.

One can, of course, consider L-functions of higher tensor powers. These L-functions contain finer information about lower rank L-functions, and it would not be an exaggeration to say that the analytic theory is still in formation.

Chapter 5
The Circle Method

1 Introduction

In his early letters to G.H. Hardy, Srinivasa Ramanujan wrote that "the coefficient of x^n in

$$\left(1 - 2x + 2x^4 - \cdots\right)^{-1}$$

is

$$\frac{1}{4n}\left(\cosh \pi\sqrt{n} - \frac{\sinh \pi\sqrt{n}}{\pi\sqrt{n}}\right) + F(\cos \pi\sqrt{n}) + f(\sin \pi\sqrt{n}).$$

I have not written here the forms of F and f as they are very irregular and complicated and their values are very small. Hence, the coefficient is an integer very near to

$$\frac{1}{4n}\left(\cosh \pi\sqrt{n} - \frac{\sinh \pi\sqrt{n}}{\pi\sqrt{n}}\right),$$

and not always the nearest integer, as I hastily wrote to you before." The significance of this passage is that Ramanujan was writing about the explicit determination of the Fourier coefficient of the reciprocal of a modular form of weight $1/2$. We see from this excerpt that he already had some intuition of the circle method even while he was in India and before he met Hardy. Indeed, in his letter, Ramanujan calculated the nth coefficient of the reciprocal of a theta series which is of weight $1/2$. In the case of the partition function, one needs to calculate the nth coefficient of the reciprocal of the Dedekind η-function, which is also a modular form of weight $1/2$. The circle method would become crystallized in their epoch-making paper [71] on the partition function written in 1918. We will make further remarks later on this point.

The idea of the method can be explained very easily. Suppose that we have a set A of numbers and we are interested in counting the number $r_{k,A}(n)$ of ways a natural number n can be written as a sum of k elements in A. It is natural to form

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_5, © Springer India 2013

the generating function

$$F_A(x) = \sum_{a \in A} x^a$$

and recognize that the coefficient of x^n in the power series expansion of $F_A(x)^k$ is precisely $r_{k,A}(n)$. However, we cannot proceed further unless we have some additional information on the function $F_A(x)$. This is often difficult and in many cases it is unreasonable to expect any further information. If, however, we look at the trigonometric series

$$f_A(\theta) = \sum_{a \in A, a \le n} e^{2\pi i a\theta},$$

then, by elementary Fourier series, we have

$$r_{k,A}(n) = \int_0^1 f_A(\theta)^k e^{-2\pi i n\theta}\, d\theta.$$

Again, we cannot proceed unless we know something about $f_A(\theta)$. But in this case, we exploit the continuity of f_A and realize that if we know the value of $f_A(a/q)$ for a rational number a/q lying in the unit interval, then we know (by partial summation) the behaviour of $f(\theta)$ in a small neighborhood of a/q. The strategy then is to break up the unit interval (the circle) into "small" intervals around rational numbers (called major arcs) where the behaviour of $f_A(\theta)$ is "known" and a complementary set of intervals (called minor arcs) and expect that the contribution from the major arcs dominates. In other words, one uses an alternate method to deal with the minor arcs to deduce that the contribution is small in comparison with the contribution from major arcs. This often is the difficult part in the application of the method. However, the method is sufficiently powerful that it enables one to make very precise conjectures on the expected behaviour of $r_{k,A}(n)$ in many interesting cases. Indeed, the behaviour of $f(a/q)$ relies on our knowledge of the distribution of the elements of A in residue classes modulo q. For example, in the case that A is the set of prime numbers, $r_{k,A}(n)$ is the number of ways of writing n as a sum of k prime numbers. Then, we have

$$f_A(a/q) = \sum_{p \le n} e^{2\pi i a p/q} = \sum_{b=0}^{q-1} e^{2\pi i a b/q} \pi(n, q, b),$$

where $\pi(n, q, b)$ is the number of primes $p \equiv b \pmod q$ with $p \le n$. From the theory of classical Dirichlet L-series, one has a great deal of information about $\pi(n, q, b)$, and this information can be injected into the formula above to derive formulas for the number of ways of writing n as a sum of k primes. For instance, in the case of the binary Goldbach problem, the major arc calculation predicts that the number of ways an even number n can be written as a sum of two prime numbers is

asymptotic to

$$2C\Big(\prod_{p|n,\,p>2}\frac{p-1}{p-2}\Big)\frac{n}{\log^2 n},$$

where C is the "twin prime constant" given by

$$\prod_{p\geq 3}\Big(1-\frac{1}{(p-1)^2}\Big),$$

and both products are over prime numbers p.

Hardy and Ramanujan developed this strategy to derive an asymptotic formula for the partition function in their epic paper of 1918. After Ramanujan's sudden demise, Hardy and Littlewood developed the method further and applied it to two classical problems, namely, Waring's problem and the Goldbach problem. In this chapter, we will discuss briefly the Hardy–Ramanujan work on the partition function which laid the foundations of the circle method, and then outline how the method can be applied to Waring's problem and the Goldbach problem. In this direction, we take an approach due to Linnik and Hua in our treatment of Waring's problem, where the circle method is needed only at the end. In the case of the Goldbach problem, we essentially follow Hardy and Littlewood in their analysis of the ternary Goldbach problem where they assumed the generalized Riemann hypothesis to get the necessary estimates on the minor arcs.

2 The Partition Function

The function $p(n)$ is the number of ways of partitioning n. For example, the number 4 can be partitioned as $1+1+1+1$, or $1+1+2$, or $2+2$, or $1+3$, or simply 4 so that $p(4)=5$. The reader should note that we do not count the orderings. Thus, the partition $1+1+2$ is the same as $2+1+1$. From the definition it is easy to see that each partition of n can be "factored" as

$$n = k_1 + 2k_2 + 3k_3 + \cdots + nk_n,$$

where k_i counts the number of times i occurs in the partition. From this observation, we deduce that $p(n)$ has the generating function given by

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{k=1}^{\infty}\big(1-x^k\big)^{-1}.$$

Why is $p(n)$ interesting? Firstly, the function occurs naturally in many places. For instance, if S_n is the symmetric group on n letters, then the conjugacy classes are in one-to-one correspondence with the partitions of n. This observation also leads

to an explicit description of the irreducible representations of S_n. Surprisingly, partitions play a fundamental role in the representation theory of general linear groups. They also have a dominant role in combinatorics and theoretical physics.

In 1918, Hardy and Ramanujan discovered, using the circle method, the asymptotic formula

$$p(n) \sim \frac{\exp(\pi\sqrt{2n/3})}{4n\sqrt{3}}$$

as n tends to infinity. In 1937, Rademacher [157] derived the following asymptotic expansion:

$$p(n) = \frac{1}{\pi\sqrt{2}} \sum_{k=1}^{\infty} \sqrt{k} A_k(n) \frac{d}{dn} \left[\frac{\sinh(\frac{\pi}{k}\sqrt{\frac{2}{3}(n-\frac{1}{24})})}{\sqrt{n-\frac{1}{24}}} \right],$$

where

$$A_k(n) = \sum_{0 \le m < k;\,(m,k)=1} e^{\pi i [s(m,k) - 2nm/k]},$$

and

$$s(b,c) = \frac{1}{4c} \sum_{n=1}^{c-1} \cot\frac{\pi n}{c} \cot\frac{\pi nb}{c}.$$

It seems that Selberg [181] had also arrived at this formula but never published it. He wrote, "In the summer of 1937, I had actually myself been studying number 36 in the Collected Papers of Ramanujan and had arrived at [Rademacher's] formula ... It always seemed strange to me that Hardy and Ramanujan did not end up with this formula ... and I believe firmly that the responsibility for this rests with Hardy."

Recently, Bruinier and Ono [28] discovered a new "exact" formula for $p(n)$. This formula expresses $p(n)$ as a finite sum of special values of a "weak" Maass form evaluated at CM points in the upper half-plane that have discriminant $-24n+1$.

To be precise, let

$$E_2(q) = 1 - 24 \sum_{n=1}^{\infty} \sigma(n) q^n, \quad \sigma(n) = \sum_{d|n} d, \quad q = e^{2\pi i z}$$

be the "quasimodular" Eisenstein series of weight two. Let $\eta(z)$ be the classical Dedekind eta function of weight $1/2$. Define

$$f(z) = \frac{E_2(z) - 2E_2(2z) - 3E_2(3z) + 6E_2(6z)}{2\eta(z)^2\eta(2z)^2\eta(3z)^2\eta(6z)^2} = \frac{1}{q} - 10 - 29q - 104q^2 \cdots.$$

One can show that $f(z)$ is a weight -2 meromorphic modular form on $\Gamma_0(6)$. Now let

$$P(z) := -\left(\frac{1}{2\pi i} \frac{d}{dz} + \frac{1}{2\pi y} \right) f(z).$$

This is a weak Maass form of weight zero. It is an eigenfunction of the hyperbolic Laplacian

$$-y^2\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)$$

with eigenvalue -2. Now consider all positive definite integral binary quadratic forms $Q(x, y) = ax^2 + bxy + cy^2$ with discriminant $b^2 - 4ac = -24n + 1$ and $6|a$. The group $\Gamma_0(6)$ acts on such forms in the obvious way, and we let Q_n be the set of these representatives with $a > 0$, $b \equiv 1 \pmod{12}$. Let

$$\alpha_Q = \frac{-b+\sqrt{1-24n}}{2a}.$$

Then, the theorem of Bruinier and Ono [29] is that

$$p(n) = \frac{1}{24n-1}\sum_{Q\in Q_n} P(\alpha_Q).$$

This formula was announced in early 2011, and upon seeing this, a natural idea that arises is to see if one can derive the Hardy–Ramanujan asymptotic for $p(n)$ from this algebraic formula. This indeed can be done, and the details will appear in a forthcoming paper of M. Dewar and M.R. Murty [42].

As noted by Hardy and Ramanujan, $p(n)$ satisfies the simple recursion formula

$$np(n) = \sum_{h=1}^{n} h \sum_{k\le n/h} p(n-hk). \tag{26}$$

To see this, we simply list all the partitions of n and count up all the numbers in this listing in two ways. The sum of all the numbers is clearly $np(n)$ which represents the left-hand side of (26). On the other hand, for each natural number h with $1 \le h \le n$, we can count how often h occurs in the listing. Indeed, the number of partitions having at least one h is clearly $p(n-h)$. Those having at least two hs is $p(n-2h)$. Thus, the number having exactly one h is

$$p(n-h) - p(n-2h).$$

Thus,

$$\begin{aligned} np(n) &= \sum_{h=1}^{n}\big\{h\big(p(n-h)-p(n-2h)\big) + 2h\big(p(n-2h)-p(n-3h)\big)+\cdots\big\} \\ &= \sum_{h=1}^{n} h \sum_{k\le n/h} p(n-hk), \end{aligned}$$

as claimed.

This simple recurrence can be used to derive a weaker result than the Hardy–Ramanujan theorem. In 1942, Erdös [48] indicated how the recursion can actually be used to derive the Hardy–Ramanujan asymptotic formula for $p(n)$. However, he was not able to evaluate the constant $1/4\sqrt{3}$. This was done later by Newman [143] in 1951. Though the proof is elementary, it is by no means simple. We refer the reader to the papers by Erdös and Newman for further details. Here, following Hardy and Ramanujan, we show that the recursion and simple induction can be used to prove that

$$\log p(n) \sim \pi\sqrt{2n/3}. \tag{27}$$

We proceed in two stages. First, we show that

$$p(n) < e^{a\sqrt{n}}, \quad a = \pi\sqrt{2/3}.$$

Second, we show that for every $\epsilon > 0$, there is a constant C_ϵ such that

$$p(n) > \frac{1}{C_\epsilon} e^{(a-\epsilon)\sqrt{n}}.$$

These two inequalities establish (27).

For the first step, we clearly have $p(1) = 1 < e^a$, and we may suppose that $p(m) < e^{a\sqrt{m}}$ has been established for $m < n$. For the recurrence (26), we have

$$np(n) < \sum_{h=1}^{n} \sum_{k=1}^{n/h} h e^{a\sqrt{n-hk}}.$$

We use the elementary inequality

$$(1-x)^u \le 1 - ux,$$

valid for any $0 < x < 1$ and $u > 0$, to deduce

$$np(n) < \sum_{h=1}^{n} \sum_{k=1}^{n/h} h e^{a\sqrt{n}(1-hk/\sqrt{n})} < e^{a\sqrt{n}} \sum_{k=1}^{n} \sum_{h=1}^{\infty} h e^{-ahk/2\sqrt{n}}.$$

Using the familiar inequality $\sinh x > x$ and

$$\sum_{m=1}^{\infty} m e^{-mx} = \frac{1}{4\sinh^2 \frac{x}{2}} < \frac{1}{x^2},$$

we deduce

$$np(n) < e^{a\sqrt{n}} \sum_{k=1}^{\infty} \frac{4n}{a^2k^2} = \frac{2\pi^2}{3a^2} n e^{a\sqrt{n}},$$

from which the estimate follows. The lower bound is derived similarly.

The central observation in the Hardy–Ramanujan paper is that the generating function

$$\sum_{n=0}^{\infty} p(n)x^n = \prod_{n=1}^{\infty} \left(1 - x^n\right)^{-1}$$

becomes, upon setting $x = e^{2\pi i\tau}$, related to the Dedekind η-function, which is a modular form of weight $1/2$.

More precisely, if

$$F(x) = \prod_{n=1}^{\infty} \left(1 - x^n\right)^{-1},$$

setting $x = e^{2\pi i\tau}$, we see that $F(x) = e^{2\pi i\tau/12}/\eta(\tau)$ where

$$\eta(\tau) = q^{1/24} \prod_{n=1}^{\infty} \left(1 - q^n\right), \quad q = e^{2\pi i\tau},$$

is the Dedekind η-function. Now, η is a modular form of weight $1/2$, and in particular,

$$\eta(-1/\tau) = \sqrt{\tau/i}\,\eta(\tau),$$

so that this translates as

$$F\left(e^{-2\pi\tau}\right) = \sqrt{\tau} e^{\pi i(1/\tau - \tau)/12} F\left(e^{-2\pi/\tau}\right).$$

In other words, if

$$\psi(\tau) = \sqrt{\tau} \exp\left(\frac{\pi}{12}\left(\frac{1}{\tau} - \tau\right)\right),$$

then as τ tends to zero, $e^{-2\pi\tau}$ tends to 1 and $e^{-2\pi/\tau}$ tends to zero, so that $F(e^{-2\pi/\tau})$ tends to 1. Thus, for x sufficiently close to 1, $F(e^{-2\pi\tau})$ is very close to $\psi(\tau)$. It is therefore reasonable to expect that for τ close to zero, $F(e^{-2\pi\tau})$ is approximated by $\psi(\tau)$. By Cauchy's theorem,

$$p(n) = \frac{1}{2\pi i} \int_C \frac{F(x)}{x^{n+1}}\, dx,$$

where C is a circle around $x = 0$ of radius < 1. Equivalently, changing variables, we have

$$p(n) = \int_{-1/2}^{1/2} F\left(e^{-2\pi\tau}\right) e^{2\pi i n\tau}\, d\tau,$$

so that if we insert our approximation for $F(e^{-2\pi\tau})$, we get

$$p(n) \sim \int_{-1/2}^{1/2} \psi(\tau) e^{-2\pi i n\tau}\, d\tau.$$

This integral is not too difficult to evaluate and is of "Bessel type." One finds that

$$p(n) \sim \frac{1}{4n\sqrt{3}} e^{a\sqrt{n}}, \quad a = \pi\sqrt{2/3}.$$

Of course, the approximation of $F(e^{2\pi\tau})$ by $\psi(\tau)$ has to be quantified. This is precisely what Hardy and Ramanujan do in their epic paper. The modular transformation property allows one to derive the asymptotic formula for $p(n)$. This is the basic strategy. The details of the paper are not simple. But the idea is.

Shortly after Ramanujan fell ill and returned to India, Hardy and Littlewood recognized the potential of the circle method in treating other problems in number theory. In a series of remarkable papers, they addressed Waring's problem and Goldbach's conjecture. We discuss these in the next two sections as they illustrate the versatility of the method.

3 Waring's Problem

In 1770, Edmund Waring, in his book "Meditationes Algebraicae" made the profound statement (without proof) that *every natural number can be written as a sum of at most four squares, nine cubes, nineteen fourth powers and so on*. More precisely, this observation, called Waring's problem, is stated as follows:

Conjecture (Waring's problem) *For each $k \geq 2$, there exists a positive integer $g = g(k) \geq 2$ such that every natural number n can be written as a sum of $g(k)$ kth powers. That is, for every $n \geq 1$, there exist non-negative integers $x_1, x_2, \ldots, x_g$ such that*

$$n = x_1^k + x_2^k + \cdots + x_g^k.$$

We note here that $g(k)$ is chosen to be the minimal number with the above property. That is, one can find at least one $n \geq 1$ which cannot be written as a sum of $(g(k) - 1)$ kth powers.

In 1909, Hilbert [80] proved that Waring's observation is true. However, his proof only showed that $g(k)$ exists but gave no method to effectively find it or bound it. Thus, we are led to the following questions.

(1) Can we find a precise formula for $g(k)$ for all k?
(2) Can we find a formula for $g(k)$ that works for *sufficiently large values of* k?
(3) Can we determine the asymptotics, lower bounds or upper bounds for $g(k)$?

In 1772, J.A. Euler, the son of L. Euler, conjectured [52] that

$$g(k) = 2^k + \left[\frac{3^k}{2^k}\right] - 2.$$

In 1936, Pillai [156] showed that Euler's conjecture is true, provided that

$$2^k\left\{\left(\frac{3}{2}\right)^k\right\}+\left[\left(\frac{3}{2}\right)^k\right]\leq 2^k.$$

By the work of Mahler [119] in 1956, it is known that there can be at most finitely many values of k for which the above inequality does not hold. It is conjectured that the above inequality holds for all k.

By the celebrated theorem of Lagrange, $g(2)=4$. In 1909, Wieferich [202] showed $g(3)=9$. A gap in his proof was filled in later by Kempner [92] in 1919. The fact that $g(4)=19$ is a recent theorem of Balasubramanian, Deshouillers and Dress [14, 15]. Earlier, Chen showed in 1964 that $g(5)=37$, and Pillai [155] in 1940 proved that $g(6)=73$. There are still many open problems in the theory. For instance, it is conjectured that every sufficiently large number can be written as the sum of four cubes.

It is possible to prove the existence of $g(k)$ by using elementary methods. This was first done by Linnik in 1943 using fundamental ideas of Schnirelmann. In this chapter, we outline Linnik's elementary solution of Waring's problem combined with a theorem of Hua on exponential sums. Then, we indicate how to use the circle method to derive an asymptotic formula for the function $r_{g,k}(n)$, the number of ways of writing n as a sum of g kth powers. In this way, the circle method comes in only at the end to enable us to derive the asymptotic formula for the number of such representations.

3.1 Schnirelmann Density

Let $A\subseteq\mathbb{N}$, and for every $n\geq 1$, let $A(n)=\#\{a\in A : a\leq n\}$. We define the *Schnirelmann density* of A, denoted $\delta(A)$ to be

$$\delta(A):=\inf_{n\geq 1}\frac{A(n)}{n}.$$

We observe that $A(n)\geq\delta(A)n$ for all $n\geq 1$. We also observe that $0\leq\delta(A)\leq 1$ and $\delta(A)=1$ if and only if $A=\mathbb{N}$.

The Schnirelmann density is different from the *asymptotic density* $\sigma(A)$ defined as

$$\sigma(A)=\lim_{n\to\infty}\frac{A(n)}{n}.$$

While $\sigma(A)$ measures the asymptotic behaviour of $\frac{A(n)}{n}$ for arbitrarily large values of n, the Schnirelmann density $\delta(A)$ is affected by the first few values of n. For example, if $\mathbb{E}$ and $\mathbb{O}$ denote the set of even and odd natural numbers respectively, then $\delta(\mathbb{E})=0$ and $\delta(\mathbb{O})=1/2$. On the other hand, $\sigma(\mathbb{E})=\sigma(\mathbb{O})=1/2$.

Given two sets A and B of integers, let $A+B$ denote the sumset of A and B, that is, $A+B=\{a+b : a\in A, b\in B\}$. If A and B are subsets of $\mathbb{N}$, let $A\oplus B$ denote the sumset of $A\cup\{0\}$ and $B\cup\{0\}$. That is,

$$A\oplus B := A\cup\{0\}+B\cup\{0\}.$$

The following remarkable theorem was proved by Schnirelmann in 1936:

Theorem 3.1 (Schnirelmann, 1936) *For any two subsets A and B of $\mathbb{N}$,*

$$\delta(A\oplus B)\geq \delta(A)+\delta(B)-\delta(A)\delta(B).$$

Proof Suppose that $A(n)=r$ and the elements $a_i\leq n$ of A are ordered as $1\leq a_1<a_2<a_3<\cdots<a_r\leq n$. Let

$$B_1 := \{b\in B : b<a_1\},$$

$$B_i := \{b\in B : a_{i-1}+b<a_i\} \quad \text{if } 2\leq i\leq r$$

and

$$B_{r+1} := \{b\in B : a_r+b\leq n\}.$$

Observe that the sets $\{a_1,a_2,\ldots,a_r\}$ and $A+B_i$ for $1\leq i\leq r+1$ are disjoint subsets of $(A\oplus B)$, and each element in these sets is $\leq n$. For notational convenience, let us define $B(0)=0$. We have

$$\begin{aligned}(A\oplus B)(n) &\geq A(n)+\sum_{i=1}^{r+1}\#B_i\\ &\geq A(n)+B(a_1-1)+\sum_{i=2}^{r}B(a_i-a_{i-1}-1)+B(n-a_r).\end{aligned}$$

Combining the above inequality with the property that $B(n)\geq\delta(B)n$, we get that for every $n\geq 1$,

$$\begin{aligned}(A\oplus B)(n) &\geq A(n)+\delta(B)\left\{(a_1-1)+\sum_{i=2}^{r}(a_i-a_{i-1}-1)+(n-a_r)\right\}\\ &= A(n)+\delta(B)(n-r)\\ &= A(n)+\delta(B)\big(n-A(n)\big)\\ &= A(n)\big(1-\delta(B)\big)+\delta(B)n\\ &\geq \delta(A)n\big(1-\delta(B)\big)+\delta(B)n\\ &= n\big(\delta(A)+\delta(B)-\delta(A)\delta(B)\big).\end{aligned}$$

This proves the theorem. □

A more general version of Schnirelmann's theorem can be stated as follows:

Theorem 3.2 *For* $A_1, A_2, \ldots, A_t \subseteq \mathbb{N}$,

$$\delta\left(\bigoplus_{i=1}^{t} A_i\right) \geq 1 - \prod_{i=1}^{t}\big(1 - \delta(A_i)\big).$$

Proof We have already proved this theorem for $t = 2$. A simple inductive exercise proves it for all $t \geq 2$. □

For any $m \in \mathbb{N}$ and $A \subseteq \mathbb{N}$, we denote

$$mA := \bigoplus_{i=1}^{m} A.$$

Lemma 3.3 *If* $\delta(B) > 1/2$ *for some* $B \subseteq \mathbb{N}$, *then* $\delta(2B) = 1$. *In other words*, $2B = \mathbb{N}$.

Proof Let $n \in \mathbb{N}$. We will show that $n \in 2B$. If $n \in B$, then we are done. If $n \notin B$, let $B_n = \{b \in B : b < n\}$ and $B'_n = \{n - b : b \in B_n\}$. Clearly, $\#B_n = \#B'_n = B(n)$, and $B_n \cup B'_n \subseteq \{1, 2, \ldots, n\}$. Thus, $\#(B_n \cup B'_n) \leq n$.

Since $\delta(B) > 1/2$, we get

$$\#B_n = B(n) = \#B'_n > \frac{n}{2}.$$

Let us assume that B_n and B'_n are disjoint. This implies $n < \#(B_n \cup B'_n)$, which contradicts the fact that $\#(B_n \cup B'_n) \leq n$. Hence, our assumption is false, and B_n and B'_n are not disjoint. In other words, there exist $b_1, b_2 < n \in B$ such that $b_1 = n - b_2$, that is, $b_1 + b_2 = n$. This proves that $n \in 2B$.

This proves the lemma. □

The following theorem of Schnirelmann connects the concept of Schnirelmann density with Waring's problem:

Theorem 3.4 *If* A *is a subset of* $\mathbb{N}$ *such that* $\delta(A) > 0$, *then there exists* $m \in \mathbb{N}$ *such that* $\delta(mA) = 1$.

Proof If $\delta(A) = 1$, we are done. If not, we have $0 < \delta(A) < 1$. Since $0 < 1 - \delta(A) < 1$, we may choose t large enough so that

$$\big(1 - \delta(A)\big)^t < \frac{1}{2}.$$

By Theorem 3.2, we see that

$$\delta(tA) \geq 1 - \big(1 - \delta(A)\big)^t > \frac{1}{2}.$$

Theorem 3.4 follows as a quick application of Lemma 3.3. □

In 1942, Mann [120] proved that Schnirelmann's theorem can be refined to show that $\delta(A \oplus B) \geq \min(\delta(A) + \delta(B), 1)$.

3.2 Schnirelmann Density and Waring's Problem

Schnirelmann introduced these ideas to attack Goldbach's conjecture and indeed, using elementary sieve methods, showed that there is a constant c_0 such that every number can be written as a sum of at most c_0 primes. Many mathematicians have successively improved upon this constant c_0, and at the moment, $c_0 = 7$, a result due to Ramaré [168]. By the same methods, one can show that every even number is the sum of at most six prime numbers. It is remarkable how far such an elementary method can be pushed.

In 1943, Linnik applied Schnirelmann's method to attack Waring's problem. We discuss this now.

Let $k \geq 2$ and $A_k = \{x^k; x \in \mathbb{N}\}$. We observe that

$$\frac{1}{n} \leq \frac{A_k(n)}{n} \leq \frac{n^{1/k}}{n} \quad \text{for every } n \geq 1.$$

Now,

$$\delta(A_k) \leq \sigma(A_k) \leq \lim_{n\to\infty} \frac{n^{1/k}}{n} = 0.$$

Thus, $\delta(A_k) = 0$. In view of Theorem 3.4 of Schnirelmann, if we could show that $\delta(mA_k) > 0$ for some $m \geq 2$, then for some $g = g(k) \geq 1$, $\bigoplus_{i=1}^{g} A_k = \mathbb{N}$. This would solve Waring's problem.

For positive integers g and m, let $r_{g,k}(m)$ denote the number of non-negative integral solutions of the equation $x_1^k + x_2^k + \cdots + x_g^k = m$. Observe that if this equation holds, then $0 \leq x_i \leq m^{1/k}$ for each $1 \leq i \leq g$. Thus,

$$r_{g,k}(m) \leq \left(m^{1/k} + 1\right)^g \ll_g m^{g/k}.$$

The following theorem of Linnik, which we will prove in the next section, shows that we can find some g for which the above estimate for $r_{g,k}(m)$ can be sharpened.

Theorem 3.5 *For any $k \in \mathbb{N}$, there exist $g \in \mathbb{N}$ and a constant $c(k)$ depending only on k such that for all $m \geq 1$,*

$$r_{g,k}(m) \leq c(k) m^{\frac{g}{k}-1}.$$

With the help of Linnik's theorem, we prove the following theorem:

Theorem 3.6 *For any $k \in \mathbb{N}$, there exists $g \in \mathbb{N}$ such that $\delta(gA_k) > 0$.*

Proof By Theorem 3.5, there exists $g \in \mathbb{N}$ such that

$$\sum_{\substack{m=0 \\ r_{g,k}(m)\neq 0}}^{n} r_{g,k}(m) \leq \sum_{\substack{m=0 \\ r_{g,k}(m)\neq 0}}^{n} c(k)m^{\frac{g}{k}-1} \leq c(k)n^{\frac{g}{k}-1} \sum_{\substack{m=0 \\ r_{g,k}(m)\neq 0}}^{n} 1. \tag{28}$$

We also observe that for each $1 \leq i \leq g$, if

$$0 \leq x_i \leq \frac{n^{1/k}}{g^{1/k}},$$

then,

$$\sum_{i=1}^{g} x_i^k \leq n.$$

Thus,

$$\sum_{\substack{x_1,x_2,\ldots,x_g \\ 0\leq x_i \leq \frac{n^{1/k}}{g^{1/k}}}} 1 \leq \sum_{\substack{m=0 \\ r_{g,k}(m)\neq 0}}^{n} r_{g,k}(m). \tag{29}$$

Observe that

$$\sum_{\substack{x_1,x_2,\ldots,x_g \\ 0\leq x_i \leq \frac{n^{1/k}}{g^{1/k}}}} 1 = \left(\left[\frac{n^{1/k}}{g^{1/k}}\right] + 1\right)^g \geq \left(\frac{n^{1/k}}{g^{1/k}}\right)^g$$

and

$$\sum_{\substack{m=0 \\ r_{g,k}(m)\neq 0}}^{n} 1 = (gA_k)(n).$$

Combining these observations with the inequalities in (28) and (29), we get

$$\left(\frac{n^{1/k}}{g^{1/k}}\right)^g \leq c(k)n^{\frac{g}{k}-1}(gA_k)(n) \quad \text{for every } n \geq 1.$$

This implies that there exists a positive constant $C(k) > 0$ such that for every $n \geq 1$,

$$(gA_k)(n) \geq C(k)n.$$

Hence, $\delta(gA_k) > 0$. □

Thus, Linnik's theorem solves Waring's problem. In the next section, we prove Linnik's theorem.

3.3 Proof of Linnik's Theorem

We start this section with a basic lemma on linear equations.

Lemma 3.7 *For a positive integer n, let $q(n)$ denote the number of integer solutions (x_1, x_2, y_1, y_2) of the equation*

$$x_1 y_1 + x_2 y_2 = n \tag{30}$$

such that $|x_i| \le X$ and $|y_i| \le Y$. Then

$$q(0) \ll (XY)^{3/2}$$

and

$$q(n) \ll \left(XY \sum_{d|n} \frac{1}{d} \right) \quad \text{for } n \ge 1.$$

Proof We first consider the case where $n = 0$. Clearly, x_1, x_2 and y_1 can take at most $2X+1$, $2X+1$ and $2Y+1$ values respectively. Once these are chosen, y_2 can take at most one value. Thus,

$$q(0) \le (2X+1)^2(2Y+1) \ll X^2 Y.$$

Similarly, $q(0) \ll XY^2$. Thus,

$$q(0) \ll \min\left\{X^2Y, XY^2\right\} \ll \sqrt{X^2Y.XY^2} \ll (XY)^{3/2}.$$

We can do better when $n \neq 0$. We assume, without loss of generality, that $X \le Y$. Let $q_1(n)$ be the number of integer solutions to $x_1 y_1 + x_2 y_2 = n$ such that $|x_2| \le |x_1| \le X$ and $|y_i| \le Y$. This ensures that $x_1 \neq 0$. Otherwise, we would get $x_2 = 0$, which implies that $n = 0$.

Let us start by fixing x_1 and x_2 and assume that $(x_1, x_2) = 1$. Given a particular solution (y_1', y_2'), all solutions of Eq. (30) are of the form

$$y_1 = y_1' + t x_2, \qquad y_2 = y_2' - t x_1, \quad t \in \mathbb{Z}.$$

We observe that

$$|t| = \frac{|y_2' - y_2|}{|x_1|} \le \frac{2Y}{|x_1|}.$$

We conclude

$$q_1(n) \le \sum_{1 \le |x_1| \le X} \sum_{|x_2| \le |x_1|} \left(2\frac{2Y}{|x_1|} + 1 \right) \le 5Y \sum_{1 \le |x_1| \le X} \frac{2|x_1| + 1}{|x_1|} \ll XY.$$

Thus, Eq. (30) has $\ll XY$ integer solutions.

If $(x_1, x_2) = d > 1$, Eq. (30) has an integer solution, provided that $d|n$. In this case, we take

$$x_1' = \frac{x_1}{d}, \qquad x_2' = \frac{x_2}{d}.$$

From above, the number of integer solutions to the equation

$$x_1' y_1 + x_2' y_2 = \frac{n}{d}$$

is $\ll XY/d$. Finally, we conclude that

$$q(n) \ll XY \sum_{d|n} \frac{1}{d}.$$

□

From the above lemma we deduce the following:

Lemma 3.8 *Let $f(x)$ be a polynomial of degree 2 with integer coefficients, say, $f(x) = c_2x^2 + c_1x + c_0$ with $c_2 = \mathcal{O}(1)$, $c_1 = \mathcal{O}(P)$ and $c_0 = \mathcal{O}(P^2)$. The number of solutions in the variables x_i and y_i such that $0 \le x_i, y_i \le P$ for $1 \le i \le 4$ to the equation*

$$f(x_1) + f(x_2) + f(x_3) + f(x_4) = f(y_1) + f(y_2) + f(y_3) + f(y_4) \tag{31}$$

is $\ll P^6$.

Proof We observe that

$$f(x_i) - f(y_i) = (x_i - y_i)\big[c_2(x_i + y_i) + c_1\big].$$

We put $z_i = x_i - y_i$ and $w_i = c_2(x_i + y_i) + c_1$.

The number of solutions of Eq. (31) is less than or equal to the number of solutions of the equation

$$z_1w_1 + z_2w_2 = -z_3w_3 - z_4w_4,$$

where $z_i \ll P$ and $w_i \ll P$. By Lemma 3.7, we see that for a fixed $n \ge 0$, the number $q(n)$ of solutions of $z_1w_1 + z_2w_2 = n$ is

$$\ll P^3 \quad \text{if } n = 0$$

and

$$\ll P^2 \sum_{d|n} \frac{1}{d} \quad \text{if } n \ge 1.$$

Thus, the number of solutions of the equation

$$z_1w_1 + z_2w_2 = -z_3w_3 - z_4w_4,$$

where $z_i \ll P$ and $w_i \ll P$, is

$$\begin{aligned}\sum_{|n|\ll P^2} q(n)^2 &\ll P^6 + \sum_{1\le n\le cP^2}\left(P^2\sum_{d|n}\frac{1}{d}\right)^2\\ &\ll P^6 + P^4\sum_{1\le n\le cP^2}\sum_{\substack{d_1|n\\ d_2|n}}\frac{1}{d_1d_2}\\ &\ll P^6 + P^4\sum_{\substack{1\le d_1\le cP^2\\ 1\le d_2\le cP^2}}\frac{1}{d_1d_2}\frac{P^2}{[d_1,d_2]}\\ &\ll P^6 + P^6\sum_{\substack{1\le d_1\le cP^2\\ 1\le d_2\le cP^2}}\frac{(d_1,d_2)}{(d_1d_2)^2}\\ &\ll P^6 + P^6\sum_{d_1=1}^{\infty}\sum_{d_2=1}^{\infty}\frac{1}{(d_1d_2)^{3/2}} \ll P^6,\end{aligned}$$

where in the penultimate step, we used the elementary observation that $(d_1,d_2)^2 \le d_1d_2$. This proves the lemma. □

Let us note that the proof of this lemma shows that the number of integer solutions of the equation

$$n_1h_1 + n_2h_2 = -n_3h_3 - n_4h_4$$

with $|n_i| \le X$ and $0 < |h_i| \le Y < X$ is $\ll (XY)^3$.

As a precursor to Linnik's theorem, we now prove the following theorem, due to Hua [83].

Theorem 3.9 *Let $k \ge 2$, and $f(x)$ be a polynomial of degree k with integer coefficients, say,*

$$f(x) = c_kx^k + c_{k-1}x^{k-1} + \cdots + c_1x + c_0$$

such that

$$c_k = \mathcal{O}(1),\ c_{k-1} = \mathcal{O}(P),\ \ldots,\ c_1 = \mathcal{O}\big(P^{k-1}\big),\ c_0 = \mathcal{O}\big(P^k\big).$$

Then

$$\int_0^1\left|\sum_{x=0}^{P} e^{2\pi i f(x)\alpha}\right|^{8^{k-1}} d\alpha = \mathcal{O}\big(P^{8^{k-1}-k}\big). \tag{32}$$

Proof Let us start with $k = 2$. Observe that

$$\int_0^1 \left| \sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right|^8 d\alpha = \int_0^1 \left(\sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right)^4 \left(\sum_{x=0}^{P} e^{-2\pi i f(x)\alpha} \right)^4 d\alpha.$$

Thus, the integral in question is equal to the number of solutions to Eq. (31), which is $\ll P^6$ by Lemma 3.8. This proves the lemma for $k = 2$. We now proceed by mathematical induction and assume that Eq. (32) holds when we replace k by $k - 1$.

Observe that

$$\begin{aligned} \left| \sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right|^2 &= \sum_{x_1=0}^{P} \sum_{x_2=0}^{P} e^{2\pi i (f(x_1)-f(x_2))\alpha} \\ &= \sum_{x=0}^{P} e^{-2\pi i f(x)\alpha} \sum_{h=-x}^{P-x} e^{2\pi i f(x+h)\alpha} \\ &= P + 1 + \sum_{h\neq 0}{}' \sum_{x}{}' e^{2\pi i (f(x+h)-f(x))\alpha}, \end{aligned}$$

where the dash on top of the summations refers to all those integers h lying between $-P$ and P and those integers x such that both $x + h$ and x lie between 0 and P. Now,

$$f(x+h) - f(x) = \sum_{j=0}^{k} c_j \left((x+h)^j - x^j \right) = \sum_{j=0}^{k} c_j \sum_{i=0}^{j-1} \binom{j}{i} x^i h^{j-i}.$$

Thus, $f(x+h) - f(x) = h\phi(x,h)$, where $\phi(x,h)$ is a polynomial in h of degree at most $k-1$. Let us define

$$a_h = \sum_{x}{}' e^{2\pi i h \phi(x,h)\alpha}.$$

Then, we have

$$\left| \sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right|^2 = P + 1 + \sum_{h\neq 0}{}' a_h.$$

Raising both sides by the power 8^{k-2}, we have

$$\left| \sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right|^{2.8^{k-2}} = \left(P + 1 + \sum_{h\neq 0}{}' a_h \right)^{8^{k-2}} \ll P^{8^{k-2}} + \left| \sum_{h\neq 0}{}' a_h \right|^{8^{k-2}}.$$

If

$$\left| \sum_{h\neq 0}{}' a_h \right| \leq P,$$

then

$$\left|\sum_{x=0}^{P} e^{2\pi i f(x)\alpha}\right|^{2.8^{k-2}} \ll P^{8^{k-2}}.$$

Hence,

$$\int_0^1 \left|\sum_{x=0}^{P} e^{2\pi i f(x)\alpha}\right|^{8^{k-1}} d\alpha \ll P^{4.8^{k-2}} \ll P^{8^{k-1}-k},$$

since $4.8^{k-2} \geq k$ for all $k \geq 2$. This proves the theorem, provided that $|\sum'_{h\neq 0} a_h| \leq P$.

Suppose now that

$$\left|\sum_{h\neq 0}' a_h\right| \geq P.$$

Then,

$$\left|P+1+\sum_{h\neq 0}' a_h\right|^{8^{k-2}} \ll \left|\sum_{0<|h|\leq P}' a_h\right|^{8^{k-2}}.$$

Applying the Cauchy–Schwarz inequality, we have

$$\left|\sum_{h\neq 0}' a_h\right|^{8^{k-2}} = \left(\left|\sum_{h\neq 0}' a_h\right|^2\right)^{2^{3(k-2)-1}} \ll \left(P\left(\sum_{h\neq 0}' |a_h|^2\right)\right)^{2^{3(k-2)-1}}.$$

Since

$$P^{2^{3(k-2)-1}+2^{3(k-2)-2}+\cdots+1}\sum_{h\neq 0}' |a_h|^{8^{k-2}} = P^{2^{3(k-2)}-1}\sum_{h\neq 0}' |a_h|^{8^{k-2}},$$

a repeated application of the Cauchy–Schwarz inequality leads to an upper bound of the form

$$\left|\sum_{h\neq 0}' a_h\right|^{8^{k-2}} \ll P^{2^{3(k-2)}-1}\sum_{h\neq 0}' |a_h|^{8^{k-2}}. \tag{33}$$

This is also immediate from an application of Hölder's inequality. So, putting it all together,

$$\begin{aligned}\left|\sum_{x=0}^{P} e^{2\pi i f(x)\alpha}\right|^{2.8^{k-2}} &= \left|P+1+\sum_{h\neq 0}' a_h\right|^{8^{k-2}} \\ &\ll \left|\sum_{0<|h|\leq P}' a_h\right|^{8^{k-2}} \ll P^{2^{3(k-2)}-1}\sum_{h\neq 0}' |a_h|^{8^{k-2}}.\end{aligned}$$

By raising Eq. (33) to the fourth power, we immediately deduce

$$\int_0^1 \left| \sum_{x=0}^{P} e^{2\pi i f(x)\alpha} \right|^{8^{k-1}} d\alpha \ll P^{4(8^{k-2}-1)} \int_0^1 \left(\sum_{h\neq 0}{}' |a_h|^{8^{k-2}} \right)^4 d\alpha. \tag{34}$$

We write

$$\left| \sum_{x=0}^{P}{}' e^{2\pi i h\phi(x,h)\alpha} \right|^{8^{k-2}} = \sum_n A(n) e^{2\pi i h n \alpha},$$

where

$$n \ll \max_{0\leq x\leq P} \left|\phi(x,h)\right| \ll P^{k-1}.$$

This gives us

$$\begin{aligned} \left|A(n)\right| &= \left| \int_0^1 \left| \sum_{x=0}^{P}{}' e^{2\pi i \phi(x,h)\alpha} \right|^{8^{k-2}} e^{-2\pi i n\alpha} d\alpha \right| \\ &\leq \int_0^1 \left| \sum_{x=0}^{P}{}' e^{2\pi i \phi(x,h)\alpha} \right|^{8^{k-2}} d\alpha, \end{aligned}$$

which by induction hypothesis is

$$\ll P^{8^{k-2}-(k-1)}.$$

Thus, by Eq. (34) and the above, we get

$$\begin{aligned} &\int_0^1 \left| \sum_{x=0}^{P}{}' e^{2\pi i f(x)\alpha} \right|^{8^{k-1}} d\alpha \\ &\ll P^{4(8^{k-2}-1)} \int_0^1 \left(\sum_{h\neq 0}{}' |a_h|^{8^{k-2}} \right)^4 d\alpha \\ &\ll P^{4(8^{k-2}-1)} \sum_{\substack{n_1,n_2,n_3,n_4 \\ h_1,h_2,h_3,h_4 \\ n_i \ll P^{k-1} \\ 0<|h_i|\leq P \\ n_1h_1+n_2h_2=-n_3h_3-n_4h_4}} A(n_1)A(n_2)A(n_3)A(n_4) \\ &\ll P^{4(8^{k-2}-1)} P^{3k} P^{4.8^{k-2}-4(k-1)} \\ &\ll P^{8^{k-1}-k}. \end{aligned}$$

This proves the theorem. □

Remark 3.1 Note that in all the above inequalities, the implied constants only depend on k.

The following corollary clearly follows from the above theorem.

Corollary 3.10 *For any $k \geq 1$, there exists $g \geq 1$ such that*

$$\int_0^1 \left| \sum_{x=0}^{P} e^{2\pi i x^k \alpha} \right|^g d\alpha \leq c(k) P^{g-k},$$

where $c(k)$ is a constant depending on k. In fact, we can take $g = 8^{k-1}$.

We are now ready to prove Theorem 3.5.
Let $k \geq 1$ and g be as in Corollary 3.10. We have

$$\left(\sum_{x=0}^{P} e^{2\pi i x^k \alpha} \right)^g = \sum_{x_1, x_2, \ldots, x_g = 0}^{P} e^{2\pi i (x_1^k + x_2^k + \cdots + x_g^k)\alpha}$$
$$= \sum_{m \geq 0} c_m e^{2\pi i m \alpha},$$

where

$$c_m = r_{g,k}(m) = \left| \int_0^1 \left(\sum_{x=0}^{P} e^{2\pi i x^k \alpha} \right)^g e^{-2\pi i m \alpha} \, d\alpha \right|.$$

By Corollary 3.10, we deduce

$$c_m \leq \int_0^1 \left| \sum_{x=0}^{P} e^{2\pi i x^k \alpha} \right|^g d\alpha \leq c(k) P^{g-k}.$$

Choosing

$$P = \left[m^{\frac{1}{k}} \right],$$

we get

$$r_{g,k}(m) \leq c(k) m^{\frac{g}{k} - 1},$$

which proves Theorem 3.5.

It is possible to derive an asymptotic formula using the circle method for the number of such representations, provided that g is large enough. Essentially, all the ingredients have been proved above with one exception, and this is the celebrated Weyl's inequality for exponential sums. Using an ingenious differencing technique, Weyl derived non-trivial estimates for exponential sums of the form

$$S := \sum_{n \leq x} e^{2\pi i \alpha P(n)},$$

where $P(n)$ is a polynomial. The idea is simple to explain. For $P(n)$ a monic linear polynomial in n, the sum is a geometric series and can easily be summed. The estimate depends on how close α is to an integer. When $P(n)$ is quadratic, then considering $|S|^2$, we get an exponential sum

$$\sum_{n,m\leq x} e^{2\pi i\alpha(P(n)-P(m))}.$$

Using the technique that already appeared in Hua's theorem, we can reduce this to the linear case and derive a non-trivial estimate for S. It is now clear that a delicate inductive process will lead to non-trivial estimates for S when P is any monic polynomial of arbitrary degree. Such an estimate was employed by Hardy and Littlewood in their treatment of Waring's problem. They were able to derive an asymptotic formula for $r_{g,k}(n)$ when $g \geq 2^k + 1$. More precisely, they showed that

$$r_{g,k}(n) = \frac{\Gamma(1+1/k)^g}{\Gamma(g/k)} n^{g/k-1}\mathfrak{S}(n) + o\big(n^{g/k-1}\big),$$

where

$$\mathfrak{S}(n) = \sum_{q=1}^{\infty} T_q(n),$$

and

$$T_q(n) = \sum_{\substack{a=1\\(a,q)=1}}^{q} q^{-g}\left(\sum_{r=1}^{q} e^{2\pi i a r^k/q}\right)^g e^{-2\pi i a n/q}.$$

The series $\mathfrak{S}(n)$ is called the *singular series* and measures the "local obstruction" to the problem of counting the number of solutions. Remarkably, $T_q(n)$ is a multiplicative function of q, and one can write the series as an infinite product from which its positivity can easily be inferred. We refer the reader to two excellent surveys on the topic for further study [47] and [197]. As can be seen from the last reference, there is still ongoing research in the development of the method.

4 Goldbach's Conjecture

In 1742, C. Goldbach conjectured that every even number greater than 2 can be written as a sum of two prime numbers. Consequently, this predicts that every odd number > 5 can be written as a sum of three primes. The binary Goldbach conjecture is still open. However, the Goldbach conjecture for odd numbers is now a theorem due to Vinogradov. His theorem is a good illustration of how one may combine elementary methods with sieve techniques to prove deep theorems. It was originally proved by Hardy and Littlewood by applying the circle method and using the generalized Riemann hypothesis. We indicate how this was done below. It is not

our intention here to give an exhaustive treatment of the method. Rather, we shall illustrate its use in the ternary Goldbach problem.

The idea can be explained as follows. Let

$$f(\theta, n) := \sum_{p \le n} e(p\theta), \qquad e(t) = e^{2\pi i t},$$

where the summation is over prime numbers. Let $r(n)$ be the number of ways of writing n as a sum of three primes. Then, clearly,

$$r(n) = \int_0^1 f^3(\theta, n) e(-n\theta)\, d\theta.$$

As explained earlier, we need to have some information on $f(\theta, n)$. When θ is a rational number, say, $\theta = a/q$, we may relate the function to primes in arithmetic progressions as follows:

$$f\left(\frac{a}{q}, n\right) = \sum_{(b,q)=1} e(ba/q)\pi(n, q, b) + O(\log q)$$

since there are at most $O(\log q)$ prime divisors of q. By the generalized Riemann hypothesis,

$$\pi(n, q, b) = \frac{\mathrm{li}(n)}{\phi(q)} + O\big(n^{1/2} \log qn\big),$$

where

$$\mathrm{li}(n) = \int_2^n \frac{dt}{\log t}.$$

Now, the sum

$$\sum_{\substack{b=1 \\ (b,q)=1}}^{q} e(ba/q)$$

is a Ramanujan sum and, as a is coprime to q, is equal to $\mu(q)$ where μ denotes the Möbius function.

Inserting this into the above expression for $f(a/q, n)$ gives

$$f\left(\frac{a}{q}, n\right) = \frac{\mu(q)}{\phi(q)} \mathrm{li}(n) + O\big(qn^{1/2} \log qn\big).$$

Thus, the GRH provides us a good approximation of $f(\theta, n)$ when $\theta = a/q$ and q is not too large. Since the function $f(\theta, n)$ is a continuous function of θ, one may apply partial summation to deduce its behaviour in a suitable neighborhood of a/q. Thus the strategy in the circle method is to decompose the interval into "major arcs" suitably controlled and the set of "minor arcs," which is just the complementary

set. I.M. Vinogradov, in his proof of the ternary Goldbach Conjecture, used the elementary sieve of Eratosthenes to estimate $f(\theta, n)$ for θ belonging to a "minor arc." Putting those two estimates together, he was able to obtain the desired asymptotic formula for $r(n)$. The use of the GRH can be replaced by the Siegel–Walfisz theorem, which is what Vinogradov did, and thus, the whole derivation can be made unconditional. As we indicate below, the method breaks down for the binary Goldbach problem.

4.1 Basic Lemmas

We recall, for the benefit of the reader, two elementary results that will be used repeatedly in the proof below. The first is a classical lemma of Dirichlet: given any positive real number θ and a rational $Q > 1$, there exist $(a, q) = 1$ with $q \le Q$ so that

$$|q\theta - a| \le \frac{1}{Q}.$$

Indeed, consider the $Q + 1$ numbers $\{n\theta\}$, $0 \le m \le Q$. If we subdivide the unit interval into Q equal parts each of length $1/Q$, we see that two of these numbers must lie in the same subinterval. Thus,

$$\big|\{m_1\theta\} - \{m_2\theta\}\big| \le \frac{1}{Q}$$

for $m_1 \ne m_2 \le Q$. Hence,

$$\big|(m_1 - m_2)\theta - \big([m_1\theta] - [m_2\theta]\big)\big| \le \frac{1}{Q}.$$

Setting $q = m_1 - m_2$ and $a = [m_1\theta] - [m_2\theta]$ gives the result.

The second elementary result is Abel's lemma. If

$$B(m) = \sum_{j \le m} b_j,$$

then,

$$\begin{aligned}\sum_{j \le m} a_j b_j &= \sum_{j \le m} a_j \big(B(j) - B(j-1)\big) \\ &= \sum_{j \le m} a_j B(j) - \sum_{j \le m-1} a_{j+1} B(j) \\ &= \sum_{j \le m} (a_j - a_{j+1}) B(j) + a_m B(m).\end{aligned}$$

Let us also recall the continuous version of partial summation:

$$\sum_{j\le m} f(j)b_j = B(m)f(m) - \int_1^m B(t)f'(t)\,dt.$$

4.2 Major Arcs

As indicated in the previous sections, we have

$$r(n) = \int_0^1 f^3(\theta, n)e(-n\theta)\,d\theta.$$

Since the integrand has period 1, we have

$$r(n) = \int_{\frac{-1}{Q}}^{1-\frac{1}{Q}} f^3(\theta, n)e(-n\theta)\,d\theta$$

for any $Q > 0$.

Now let $(a, q) = 1$, $a \le q$, $q \le \log^c n$, $Q = n/(\log^c n)$, where c is a constant to be chosen.

Define $I(a, q)$ as the set of points θ such that

$$\left|\theta - \frac{a}{q}\right| \le \frac{1}{Q}.$$

Observe that no two distinct intervals $I(a, q)$ overlap for if $\theta \in I(a, q)$ and $I(a', q')$, then because $aq' \ne a'q$, we have

$$\frac{1}{qq'} \le \left|\frac{a}{q} - \frac{a'}{q}\right| = \left|\left(\frac{a}{q} - \theta\right) + \left(\theta - \frac{a'}{q'}\right)\right| \le \frac{2}{Q},$$

which implies $Q \le 2qq' \le 2\log^{2c} n$, a contradiction for n large enough. The major arcs are by the definition

$$\mathfrak{M} = \bigcup_{q\le \log^c n} \bigcup_{(a,q)=1} I(a, q).$$

The minor arcs are $[0, 1]\backslash\mathfrak{M}$, or more accurately

$$\mathfrak{m} = \left[\frac{-1}{Q}, 1 - \frac{1}{Q}\right]\backslash\mathfrak{M}.$$

We then have

$$r(n) = \int_{\mathfrak{M}} f^3(\theta, n)e(-n\theta)\,d\theta + \int_{\mathfrak{m}} f^3(\theta, n)e(-n\theta)\,d\theta,$$

and we analyze each of these intervals separately.

4.3 Application of Partial Summation

Let us begin with the observation that for any trigonometric sum of the form

$$S(\theta, n) = \sum_{m \le n} a_m e(m\theta),$$

we have by partial summation

$$S(\theta_1 + \theta_2, n) = S(\theta_1, n)e(\theta_2 n) - 2\pi i\theta_2 \int_0^n S(\theta_1, u)e(\theta u)\, du.$$

Applying this to $f(\theta, n)$ gives

$$f(\theta_1 + \theta_2, n) = e(\theta_2 n) f(\theta_1, n) - 2\pi i\theta_2 \int_0^n f(\theta_1, u)e(\theta_2 u)\, du.$$

We also introduce the function

$$g(\theta, n) = \sum_{m \le n} \frac{1}{\log m} e(m\theta)$$

and find

$$g(\theta_1 + \theta_2, n) = e(\theta_2 n) g(\theta_1, n) - 2\pi i\theta_2 \int_0^n g(\theta_1, u)e(\theta_2 u)\, du.$$

4.4 Primes in Arithmetic Progressions

Applying the Siegel–Walfisz theorem, we find when $\theta = a/q$,

$$f\left(\frac{a}{q}, n\right) = \frac{\mu(q)}{\phi(q)} \mathrm{li}(n) + O\left(\frac{n}{\log^A n}\right),$$

where we have used the fact that

$$\mu(q) = \sum_{\substack{b=1 \\ (b,q)=1}}^{q} e(ba/q),$$

a familiar Ramanujan sum. Let us observe that

$$\begin{aligned} g(0, n) &= \sum_{m \le n} \frac{1}{\log m} = \int_2^n \frac{dt}{\log t} + O\left(\frac{1}{\log n}\right) \\ &= \mathrm{li}(n) + O\left(\frac{1}{\log n}\right). \end{aligned}$$

Thus,

$$f\left(\frac{a}{q}, n\right) = \frac{\mu(q)}{\phi(q)} g(0,n) + O\left(\frac{n}{\log^A n}\right).$$

By partial summation,

$$f\left(\frac{a}{q} + y, n\right) = e(yn) f\left(\frac{a}{q}, n\right) - 2\pi i y \int_0^n f\left(\frac{a}{q}, u\right) e(yu)\, du$$

and

$$g(y,n) = e(yn) g(0,n) - 2\pi i y \int_0^n g(0,u) e(yu)\, du,$$

so that

$$\begin{aligned}
&\left| f\left(\frac{a}{q} + y, n\right) - \frac{\mu(q)}{\phi(q)} g(y,u) \right| \\
&\quad \le \frac{n}{\log^A n} + 2\pi |y| \int_0^n \left| f\left(\frac{a}{q}, u\right) - \frac{\mu(q)}{\phi(q)} g(0,u) \right| du \\
&\quad \ll \frac{n}{\log^A n} \quad \text{if } |y| \le \frac{\log^A n}{n}.
\end{aligned}$$

Since

$$\left|z^3 - w^3\right| = |z - w| \left|z^2 + w^2 + zw\right|,$$

we immediately deduce

$$\left| f^3\left(\frac{a}{q} + y, n\right) - \frac{\mu(q)}{\phi(q)^3} g^3(y,n) \right| \le \frac{3n^3}{\log^A n} \qquad (*)$$

for $|y| \le (\log^A n)/n$.

4.5 The Singular Series

We can now write down the main term of the asymptotic formula. Let us put

$$J(a,q) := \int_{I(a,q)} f^3(\theta, n) e(-n\theta)\, d\theta$$

and

$$J_1 = \int_{-1/Q}^{1/Q} g^3(y,n) e(-ny)\, dy,$$

so that the inequality of the previous section leads to

$$\left| J(a,q) - \frac{\mu(q)}{\phi^3(q)} J_1 e\left(-\frac{na}{q}\right)\right| \le \frac{3n^3}{\log^A n}\frac{\log^c n}{n} \le \frac{3n^2}{\log^{A-c} n}. \qquad (**)$$

We can choose A as large as we please. If we put

$$\rho(n) = \sum_{m_1+m_2+m_3=n, m_i\ge 2} (\log m_1)^{-1}(\log m_2)^{-1}(\log m_3)^{-1},$$

then

$$\rho(n) = \int_{-1/2}^{1/2} g^3(y,n) e(-ny)\, dy.$$

Let us note that

$$\frac{1}{3}\frac{n^2}{\log^3 n} \le \rho(n) \le n^2,$$

as is easily verified. Moreover,

$$\left|\sum_{m\le n} e(my)\right| \ll \frac{1}{|y|}$$

as is also easily deduced since the sum is a geometric series. Thus, by Abel's lemma,

$$g(y,n) = \sum_{m\le n} \frac{e(my)}{\log m} \ll \frac{1}{|y|}.$$

Thus,

$$\left|\int_{1/Q}^{1/2} g^3(y,n) e(-ny)\, dy\right| \ll Q^2 \ll \frac{n^2}{\log^{2c} n}.$$

A similar estimate holds for the integral over the range $[-\frac{1}{2}, -\frac{1}{Q}]$. Thus, by $(**)$,

$$\rho(n) = \int_{-1/Q}^{1/Q} g^3(y,n) e(-ny)\, dy + O\left(\frac{n^2}{\log^{2c} n}\right).$$

Hence, from $(*)$ we immediately see that

$$\left| J(a,q) - \frac{\mu(q)}{\phi^3(q)}\rho(n) e\left(-\frac{na}{q}\right)\right| \ll \frac{n^2}{\log^{A-c} n}.$$

We denote the Ramanujan sum

$$c_q(n) = \sum_{\substack{a=1\\(a,q)=1}}^{q} e\left(-\frac{na}{q}\right),$$

so that

$$\left| \sum_{q \le \log^c n} \sum_{(a,q)=1} J(a,q) - \rho(n) \sum_{q \le \log^c n} \frac{\mu(q)}{\phi^3(q)} c_q(n) \right|$$
$$\ll \frac{n^2}{\log^{A-c} n} \log^{2c} n \ll \frac{n^2}{\log^{A-3c} n}.$$

Thus,

$$\int_{\mathfrak{M}} f^3(\theta, n) e(-\theta n)\, d\theta = \rho(n) \sum_{q \le \log^c n} \frac{\mu(q)}{\phi^3(n)} c_q(n) + O\left(\frac{n^2}{\log^{A-3c} n} \right),$$

and the sum on the right-hand side is called the *singular series*. It is not difficult to see that

$$\sum_{q \le \log^c n} \frac{\mu(q)}{\phi^3(q)} c_q(n) = \sum_{q=1}^{\infty} \frac{\mu(q)}{\phi^3(q)} c_q(n) + O\left(\frac{1}{\log^c n} \right),$$

and using the fact that $c_q(n)$ is multiplicative, we find that the series is

$$\prod_p \left(1 - \frac{c_p(n)}{(p-1)^3} \right).$$

When n is odd, this series is

$$2 \prod_{p>2} \left(1 - \frac{c_p(n)}{(p-1)^3} \right) \ge 2 \prod_{p>2} \left(1 - \frac{1}{(p-1)^2} \right) \ge 2 \prod_{m=1}^{\infty} \left(1 - \frac{1}{m^2} \right) = 1.$$

Thus, in particular, the series is not zero. If we can show

$$\left| \int_{\mathfrak{m}} f^3(\theta, n) e(-n\theta)\, d\theta \right| \ll \frac{n^2}{\log^4 n},$$

we can conclude that

$$r(n) = \rho(n) \prod_p \left(1 - \frac{c_p(n)}{(p-1)^3} \right) + O\left(\frac{n^2}{\log^4 n} \right)$$

and that

$$r(n) \gg \frac{n^2}{\log^3 n},$$

completing the proof.

4.6 The Minor Arcs Estimate Using GRH

We begin by noting that

$$\left|\int_{\mathfrak{m}} f^3(\theta,n)e(-n\theta)\,d\theta\right| \le \sup_{\theta\in\mathfrak{m}} \left|f(\theta,n)\right| \int_0^1 \left|f^2(\theta,n)\right| d\theta \le \pi(n) \sup_{\theta\in\mathfrak{m}} \left|f(\theta,n)\right|$$

by Parseval's formula. Following Hardy and Littlewood, we want to indicate how the generalized Riemann hypothesis (GRH) can be used to supply such an estimate in a quick way.

Let us observe that (as before)

$$\begin{aligned} f\left(\frac{a}{q},n\right) &= \sum_{p\le n} e(pa/q) \\ &= \sum_{(b,q)=1} e(ba/q)\pi(n,q,b) \\ &= \frac{\mu(q)}{\phi(q)}\mathrm{li}(n) + \sum_{(b,q)=1} e(ba/q)E(n,q,b), \end{aligned}$$

where

$$E(n,q,b) = \pi(n,q,b) - \frac{\mathrm{li}(n)}{\phi(q)}.$$

Now,

$$E(n,q,b) = \frac{1}{\phi(q)} \sum_{\chi\ne\chi_0} \overline{\chi}(b)\pi(n,\chi),$$

where the sum is over non-trivial characters $\chi \pmod q$. Thus,

$$\sum_{(b,q)=1} e(ba/q)E(n,q,b) = \frac{1}{\phi(q)} \sum_{\chi\ne\chi_0} \chi(a)\tau(\overline{\chi})\pi(n,\chi),$$

where $\tau(\chi)$ denotes the Gauss sum attached to χ. The GRH is equivalent to the estimate

$$\left|\pi(n,\chi)\right| \ll n^{1/2}\log qn,$$

so that we get

$$f\left(\frac{a}{q},n\right) = \frac{\mu(q)}{\phi(q)}\mathrm{li}(n) + O\left(q^{1/2}n^{1/2}\log qn\right).$$

With these remarks, we may proceed to estimate $f(\theta,n)$ when θ is an element of the minor arcs. We observe that Dirichlet's lemma gives that there are q,a such that

$|q\theta - a| \le 1/Q$. As $\theta \in \mathfrak{m}$, we must have $q \ge \log^c n$. Thus,

$$\left|\theta - \frac{a}{q}\right| \le \frac{1}{qQ} \le \frac{1}{n}.$$

Let $y = \theta - a/q$. Then, by partial summation,

$$f(\theta, n) = f\left(\frac{a}{q} + y, n\right) = e(yn) f\left(\frac{a}{q}, n\right) - 2\pi i y \int_0^n f\left(\frac{a}{q}, u\right) e(yu)\, du,$$

so that

$$\left|f(\theta, n)\right| \ll \frac{n}{\log^{c-1} n}.$$

Choosing c sufficiently large gives us the desired estimate for the minor arcs.

The elimination of the use of GRH by Vinogradov emanates from two simple ideas. The first is an estimate for a trigonometric sum. The second is a clever application of the sieve of Eratosthenes. Combining both of these ideas, Vinogradov showed that the estimate of the function $f(\theta, n)$ on the minor arcs can be deduced without GRH. We refer the reader to [197] for details.

The binary Goldbach problem seems to lie at a deeper level. One can derive the expected main term by analyzing the major arcs as we have done in the ternary Goldbach problem. However, the minor arc estimates are not good enough to yield the desired result. Clearly, a new idea needs to be injected into the method.

These two examples, showing how Waring's problem and the ternary Goldbach problem can be attacked using the circle method, should illustrate the power and scope of the method introduced by Hardy and Ramanujan in their foundational paper giving the asymptotics of the partition function. The method is still a sophisticated tool in additive analytic number theory and very much the focus of current research.

Chapter 6
Ramanujan and Transcendence

1 Nesterenko's Theorems

Although Ramanujan did not work in transcendental number theory, his intuition in the theory of q-series, singular moduli, and the theory of quasi-modular forms played a pivotal role in its development. Indeed, his celebrated paper of 1916 entitled "On certain arithmetical functions" is notable for its wealth of ideas, not only for its enunciation of the Ramanujan conjecture regarding the τ-function, but also for its emphasis on Eisenstein series and the differential equations satisfied by them. More precisely, Ramanujan considers the functions

$$P(q) = 1 - 24\sum_{n=1}^{\infty}\sigma_1(n)q^n,$$

$$Q(q) = 1 + 240\sum_{n=1}^{\infty}\sigma_3(n)q^n,$$

$$R(q) = 1 - 504\sum_{n=1}^{\infty}\sigma_5(n)q^n,$$

where

$$\sigma_k(n) = \sum_{d|n} d^k.$$

These functions are connected to the classical Eisenstein series $E_2(z)$, $E_4(z)$, $E_6(z)$ via the equations

$$E_2(z) = P\big(e^{2\pi iz}\big), \qquad E_4(z) = Q\big(e^{2\pi iz}\big), \qquad E_6(z) = R\big(e^{2\pi iz}\big)$$

with $z \in \mathbb{C}$ and $\Im(z) > 0$. E_4 and E_6 are modular forms of weight 4 and 6 respectively, whereas E_2 is a quasi-modular form of weight 2. It is known [207] that the

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_6, © Springer India 2013

algebra of quasi-modular forms is closed under the differential operator $D = qd/dq$ and is generated by E_2, E_4, and E_6.

In his 1916 paper, Ramanujan shows that

$$DP = \frac{1}{12}\big(P^2 - Q\big), \qquad DQ = \frac{1}{3}(PQ - R), \qquad DR = \frac{1}{2}\big(PR - Q^2\big).$$

These equations played a central role in the remarkable theorem of Nesterenko [142] concerning the special values of these functions.

Theorem 1.1 (Nesterenko, 1996) *For any complex number q with $0 < |q| < 1$, the set*

$$\big\{q, P(q), Q(q), R(q)\big\}$$

contains at least three numbers algebraically independent over $\mathbb{Q}$.

As Ramanujan was aware, the functions P, Q, R are related to many interesting elliptic and modular functions. This allows one to relate the above theorem to several specific constants. For instance, if we take $q = e^{-2\pi}$, then one can show that

$$P\big(e^{-2\pi}\big) = \frac{3}{\pi}, \qquad Q\big(e^{-2\pi}\big) = \frac{3\Gamma(1/4)^8}{(2\pi)^6}, \qquad R\big(e^{-2\pi}\big) = 0.$$

Thus, Nesterenko's theorem implies the following:

Corollary 1.2 *The numbers π, e^{π}, and $\Gamma(1/4)$ are algebraically independent over $\mathbb{Q}$.*

More generally, let $\wp(z)$ be the Weierstrass $\wp$-function with invariants g_2, g_3. Let the corresponding lattice be generated by ω_1, ω_2 and denote by η_1, η_2 the associated quasi-periods. We suppose that $z = \omega_2/\omega_1$ with $\Im(z) > 0$. By the theory of elliptic functions (see Chap. 4 of [107]), we have upon setting $q = e^{2\pi i z}$,

$$P(q) = \frac{3\omega_1\eta_1}{\pi^2}, \qquad Q(q) = \frac{3}{4}\left(\frac{\omega_1}{\pi}\right)^4 g_2, \qquad R(q) = \frac{27}{8}\left(\frac{\omega_1}{\pi}\right)^6 g_3.$$

These formulas imply that $P(q), Q(q), R(q)$ belong to the field $\mathbb{Q}(g_2, g_3, \omega_1/\pi, \eta_1/\pi)$. From Nesterenko's theorem, we have

Corollary 1.3 *Let $\wp(z)$ be the Weierstrass $\wp$-function with algebraic invariants g_2, g_3. Then the numbers*

$$e^{2\pi i(\omega_2/\omega_1)}, \quad \omega_1/\pi, \quad \eta_1/\pi$$

are algebraically independent over $\mathbb{Q}$.

If the elliptic curve associated with $\wp(z)$ has complex multiplication, we deduce the following:

Corollary 1.4 *Let $\wp(z)$ be the Weierstrass $\wp$-function with algebraic invariants g_2, g_3 and whose associated elliptic curve has complex multiplication by the imaginary quadratic field k. If ω is any period of $\wp(z)$, and η is the corresponding quasi-period and $z \in k$ with $\Im(z) \neq 0$, then both sets*

$$\left\{\pi, \omega, e^{2\pi i z}\right\} \quad \text{and} \quad \left\{\omega, \eta, e^{2\pi i z}\right\}$$

are algebraically independent over $\mathbb{Q}$.

Indeed, in the case of complex multiplication, ω_2, η_2 are algebraic over the field $\mathbb{Q}(\omega_1, \eta_1)$ (see Chap. 3 of [122]). Using the Legendre relation

$$\omega_2\eta_1 - \omega_1\eta_2 = 2\pi i,$$

we deduce that η_1 is algebraic over $\mathbb{Q}(\omega_1, \pi)$ and that π is algebraic over $\mathbb{Q}(\omega_1, \eta_1)$. This leads to the above corollary.

For any natural number d, there exists a Weierstrass $\wp$-function with algebraic invariants and complex multiplication by $\mathbb{Q}(\sqrt{-d})$. Thus, the previous corollary implies the following:

Corollary 1.5 *For any natural number d, the numbers*

$$\pi, \quad e^{\pi\sqrt{d}}$$

are algebraically independent over $\mathbb{Q}$.

In fact, a stronger result can be deduced from the above. Using the Chowla–Selberg formula [34] and its relation to the theory of elliptic curves [61], we find the following:

Corollary 1.6 *For any imaginary quadratic field with discriminant $-D$ and quadratic character ϵ, the numbers*

$$\pi, \quad e^{\pi\sqrt{D}}, \quad \prod_{a=1}^{D} \Gamma(a/D)^{\epsilon(a)}$$

are algebraically independent over $\mathbb{Q}$.

In the special case that $D = 3$, we deduce, from the corollary upon using the functional equation $\Gamma(z)\Gamma(1-z) = \pi/\sin\pi z$ with $z = 1/3$, the following:

Corollary 1.7 *The numbers π, $e^{\pi\sqrt{3}}$, and $\Gamma(1/3)$ are algebraically independent.*

2 Special Values of the Γ-Function at CM Points

It is, in fact, possible to deduce further transcendence results from Nesterenko's theorem, in the context of Ramanujan's work. In 1915, Ramanujan [158] wrote a paper evaluating the product

$$\prod_{n=0}^{\infty}\left(1+\left(\frac{x}{a+nd}\right)^3\right).$$

In this paper, we find some elegant formulas of the following kind:

$$\prod_{n=1}^{\infty}\left(1+\left(\frac{2\alpha}{\alpha+n}\right)^3\right)=\frac{\Gamma(1+\alpha)^3\sinh(\pi\alpha\sqrt{3})}{\Gamma(1+3\alpha)\pi\alpha\sqrt{3}}$$

and

$$\prod_{n=1}^{\infty}\left(1+\left(\frac{2\alpha+1}{\alpha+n}\right)^3\right)=\frac{\Gamma(1+\alpha)^3\cosh(\pi(\frac{1}{2}+\alpha)\sqrt{3})}{\Gamma(2+3\alpha)\pi}.$$

If we put $\alpha=0$ in the second formula, we deduce the strikingly beautiful evaluation

$$\prod_{n=1}^{\infty}\left(1+\frac{1}{n^3}\right)=\frac{\cosh(\pi\sqrt{3}/2)}{\pi}, \tag{35}$$

which is transcendental by Nesterenko's theorem. One would expect these products to be transcendental for every rational α, but this deduction can only be made (at present) for a limited number of rational values of α like $\alpha=1/3$, $1/4$, or $1/6$. Other products appear in Chaps. 13 and 14 of Ramanujan's notebooks. For instance, in [159] we find

$$\prod_{n=1}^{\infty}\left(1+\left(\frac{x}{\alpha+n}\right)^2\right)=\frac{|\Gamma(\alpha)|^2}{|\Gamma(\alpha+ix)|^2}. \tag{36}$$

Since

$$\frac{\sin\pi z}{\pi z}=\prod_{n=1}^{\infty}\left(1-\frac{z^2}{n^2}\right),$$

we immediately deduce from Ramanujan's formula (36) and Nesterenko's theorem the following:

Theorem 2.1 *Let $a+b\sqrt{-D}$ with $a\in\mathbb{Z}$ and $b\in\mathbb{Q}$ be a CM point. Then, $\Gamma(a+b\sqrt{-D})$ is transcendental. In particular, $\Gamma(ix)$ is transcendental for every non-zero rational x.*

We refer to the forthcoming paper [138] for more applications of these ideas in the context of infinite products.

3 Special Values of Jacobi's Theta Series

In Part V of Bruce Berndt's epic series on Ramanujan's notebooks, we find Chapt. 35 dedicated to special values of the theta functions (in Ramanujan's notation)

$$\varphi(q) = \sum_{-\infty}^{\infty} q^{n^2}$$

and

$$\psi(q) = \sum_{n=0}^{\infty} q^{n(n+1)/2}.$$

From the Jacobi triple product identity it is not hard to see that these series can be written as infinite products. More precisely, we have [44]

$$\varphi(q) = \frac{\eta(q^2)^5}{\eta(q)^2\eta(q^4)}$$

and

$$\psi(q) = \frac{\eta(q^4)^2}{\eta(q^2)},$$

where η denotes the Dedekind η-function.

In his notebooks, Ramanujan evaluated many special values of φ and ψ. All of his results are related to either $\varphi(e^{-\pi x})$ or $\psi(e^{-\pi x})$ for certain rational values of x. The striking feature of his evaluations, tabulated on p. 325 of [20], is that the first group ((i)–(x)) of values of $\varphi(q)$ are algebraic multiples of

$$a = \frac{\pi^{1/4}}{\Gamma(3/4)},$$

whereas the second group of values of $\psi(q)$ are algebraic multiples of $ae^{r\pi/32}$ where r is a natural number.

The chapter in [20] provides no theoretical explanation for these formulas. Nor does it give any reason why only $\pi^{1/4}$, $\Gamma(3/4)$, and e^{π} appear as the transcendental quantities. This fact could have only been deduced from the work of Nesterenko.

We indicate briefly the theoretical rationale for these results. The detailed calculations will be given in a forthcoming paper. For the moment, suffice it to say that the 24th power of $\varphi(q)$ or $\psi(q)$ allows us to express it in terms of Ramanujan's Δ-function. This, in turn, can be expressed in terms of $E_4(z)$ and $E_6(z)$, and then Nesterenko's work can be invoked.

4 The Rogers–Ramanujan Continued Fraction

The Rogers–Ramanujan continued fraction is

$$F(q) = \cfrac{q^{1/5}}{1 + \cfrac{q}{1 + \cfrac{q^2}{1 + \ddots}}}.$$

Its genesis is rooted in the celebrated Rogers–Ramanujan identities. These are remarkable identities expressed by the following equations:

$$G(q) := \sum_{m=0}^{\infty} \frac{q^{m^2}}{(1-q)(1-q^2)\cdots(1-q^m)} = \prod_{m=1}^{\infty} \left(1-q^{5m+1}\right)^{-1} \left(1-q^{5m+4}\right)^{-1},$$

$$H(q) := \sum_{m=0}^{\infty} \frac{q^{m(m+1)}}{(1-q)(1-q^2)\cdots(1-q^m)} = \prod_{m=1}^{\infty} \left(1-q^{5m+2}\right)^{-1} \left(1-q^{5m+3}\right)^{-1}.$$

The main feature of these identities is their striking beauty and simplicity. In 1936, Selberg [178] found generalizations of these identities. Closer examination of these identities reveals hidden modular connections.

The Rogers–Ramanujan identities have an exotic history. Apparently, they were first discovered by Rogers [175] in 1894. It seems that he was a gifted mathematician whose work was largely ignored. For instance, he discovered Hölder's inequality before Hölder did (see p. 100 of [67]). The identities were re-discovered by Ramanujan in 1913, before he had come to England. However, he had no proof and knew he had none. Since no one in Hardy's circle of mathematicians could prove them, they were stated without proof in the second volume of MacMahon's book on combinatorial analysis. In 1917, Ramanujan was looking through some old volumes of the Proceedings of the London Mathematical Society and stumbled on the 1894 paper of Rogers. Hardy wrote, "I can remember very well his surprise, and the admiration which he expressed for Rogers's work. A correspondence followed in the course of which Rogers was led to a considerable simplification of his original proof." Around the same time, Schur [177] rediscovered these identities and published two combinatorial proofs. In 1919, Ramanujan [163] published another (more analytic) proof.

In his 1894 paper, Rogers [175] proved that

$$q^{-1/5} F(q) = \frac{H(q)}{G(q)}.$$

Using the Rogers–Ramanujan identities, the right-hand side can be written as

$$\prod_{n=1}^{\infty} (1-q^n)^{\chi_5(n)},$$

where $\chi_5(n)$ is the non-trivial quadratic character (mod 5) given by the Legendre symbol $(5/n)$. If we put $q = e^{2\pi iz}$, then writing the right-hand side in terms of theta functions, one can show that in fact, $F(e^{2\pi iz})$ is a modular function for $\Gamma(5)$ (see, for example, [43]).

The stage is now set to apply Nesterenko's theorems to $F(q)$. These were derived by Duverney, Ke. Nishioka, Ku. Nishioka, and Shiokawa [44]. The explicit relationship of this function to modular functions is given by the following equation:

$$\frac{1}{F(q)} = F(q) + 1 + q^{2/5}\frac{\eta(q^{1/5})}{\eta(q^5)},$$

where η is the Dedekind η-function. In [44], it was shown that Nesterenko's theorem implies the following:

Theorem 4.1 *$F(\alpha)$ is transcendental for any algebraic α with $0 < |\alpha| < 1$.*

In Ramanujan's letters to Hardy, we find several explicit evaluations of F. For instance, in his first letter, Ramanujan wrote

$$F\left(e^{-2\pi}\right) = \sqrt{\frac{5+\sqrt{5}}{2}} - \frac{1+\sqrt{5}}{2}.$$

He further wrote that $F(e^{-\pi\sqrt{n}})$ can be determined exactly for any natural number n. From the modular perspective, this is now clear since F being a modular function means that it is a rational function in $j(z)$ and $j(5z)$, where j is the j-function. Ramanujan was evaluating the function F at CM points in the upper half-plane, and for such values, the j-function assumes algebraic values by a classical theorem of Schneider (see, for example, [135]).

5 Nesterenko's Conjectures

Nesterenko conjectures that if $z \in \mathbb{C}$, $\Im(z) > 0$, and $q = e^{2\pi iz}$, then the following is true. If the set

$$\left\{z, q, P(q), Q(q), R(q)\right\}$$

contains at most three algebraically independent numbers over $\mathbb{Q}$, then z lies in an imaginary quadratic field, and the numbers

$$q, P(q), Q(q)$$

are algebraically independent.

The work of Nesterenko allows us to study transcendental values of modular forms at algebraic arguments. An example of this is illustrated by the following theorem proved by Gun, Murty, and Rath [64]. To state the theorem, we say that

two modular forms f, g are equivalent if there are two natural numbers k_1, k_2 such that $f^{k_2} = \lambda g^{k_1}$, with λ algebraic.

Theorem 5.1 *Let α be a non-CM point in the upper half-plane. Let S_α be the set of non-zero modular forms with algebraic coefficients, of arbitrary weight for the full modular group, for which $f(\alpha)$ is algebraic. Then up to equivalence, S_α has at most one element.*

If we assume Nesterenko's conjecture, then there is no exception, and the set S_α is empty.

6 Special Values of the Riemann Zeta Function and q-Analogues

In Entry 21(i) of Chap. 14 in Ramanujan's second notebook, we find the following remarkable formula. If α, β are positive numbers such that $\alpha\beta = \pi^2$, and if r is a positive integer, then

$$\begin{aligned}(4\alpha)^{-r}&\left(\frac{1}{2}\zeta(2r+1) + \sum_{m=1}^{\infty}\frac{1}{m^{2r+1}(e^{2m\alpha}-1)}\right)\\ &- (-4\beta)^{-r}\left(\frac{1}{2}\zeta(2r+1) + \sum_{m=1}^{\infty}\frac{1}{m^{2r+1}(e^{2m\beta}-1)}\right)\\ &= -\sum_{k=0}^{r+1}\frac{(-1)^k B_{2k} B_{2r+2-2k}\alpha^{r+1-k}\beta^k}{(2k)!(2r+2-2k)!},\end{aligned}$$

where the B_js are the Bernoulli numbers. (See pp. 275–276 of [19].) Apart from its intrinsic beauty, this formula is interesting for a variety of reasons. We know that the special value of the Riemann zeta function $\zeta(s)$ for $s = 2k$ is a rational multiple of π^{2k}. If $s = 2k+1$ is odd, the transcendence of $\zeta(2k+1)$ has not been established. Ramanujan's formula expresses these values as a rational multiple of π^{2k+1} plus an additional term which can be interpreted as an Eichler integral. The transcendence of this Eichler integral is studied in a recent paper of Gun, Murty, and Rath [65]. We elaborate on this below.

Apparently, the formula also appears in Ramanujan's lost notebooks, and from this we can surmise that it was discovered before Ramanujan's arrival in England in 1914. In [22], Berndt relates that several mathematicians re-discovered special cases of this formula, the most notable being Emil Grosswald [62], who extended its range of applicability to the upper half-plane. For instance, in the special case $r = 1$, we get

$$\zeta(3) + 2\sum_{n=1}^{\infty}\frac{1}{n^3(e^{2\pi n}-1)} = \frac{7\pi^3}{180}.$$

Thus, at least one of the terms on the left-hand side is transcendental! Similarly, one can deduce that at least one of

$$\zeta(4k+3), \quad \sum_{n=1}^{\infty} \frac{1}{n^{4k+3}(e^{2\pi n}-1)}$$

is transcendental for every integer $k \geq 0$. These identities suggest we consider the function

$$F_k(z) = \sum_{n=1}^{\infty} \sigma_{-k}(n) e^{2\pi i n z}.$$

One can re-write this as

$$-\zeta(k) - \sum_{n=1}^{\infty} \frac{1}{n^k(e^{2\pi i n z}-1)}.$$

For odd negative values of $k \leq -3$, $F_k(z)$ is essentially the Eisenstein series $E_{1-k}(z)$ (apart from the constant term) of weight $1-k$ for the full modular group. For odd positive values of k, one can view $F_k(z)$ as an Eichler integral, which is obtained by integrating a classical modular form. Indeed, noting that

$$\sigma_{-k}(n) = n^{-k}\sigma_k(n),$$

we may re-write $F_k(z)$ as

$$\sum_{n=1}^{\infty} \frac{\sigma_k(n)}{n^k} e^{2\pi i n z}.$$

It is now clear that $F_k(z)$ can be obtained by successive integrations of the classical Eisenstein series minus its constant term. $F_k(z)$ is not quite a modular form, and it seems natural to study what happens to it under modular transformations. Thus, for any given $k \geq 4$, we can look at

$$F_{2k+1}(z) - z^{2k} F_{2k+1}(-1/z)$$

and study the special values of this function at algebraic points. The reason for doing this is clear since this expression is, by the Ramanujan–Grosswald formula,

$$\frac{1}{2}\zeta(2k+1)\big(z^{2k}-1\big) + \frac{(2\pi i)^{2k+1}}{2z} \sum_{j=0}^{k+1} z^{2k+2-2j} \frac{B_{2j} B_{2k+2-2j}}{(2j)!(2k+2-2j)!}.$$

If for some value of z which is not a $2k$th root of unity, we have a zero of the polynomial appearing in the summation, we can then write $\zeta(2k+1)$ as a difference

of two Eichler integrals. Indeed, the zeros of the polynomial

$$R_{2k+1}(z) = \sum_{j=0}^{k+1} z^{2k+2-2j} \frac{B_{2j} B_{2k+2-2j}}{(2j)!(2k+2-2j)!},$$

called the Ramanujan polynomial (of degree $2k+2$) by the authors in [137], has been studied by Murty, Smyth, and Wang. For $k \geq 4$, they showed that all the zeros of $R_{2k+1}(z)$ are simple and lie on the unit circle, apart from four real roots. Moreover, the only possible roots of unity which are zeros of $R_{2k+1}(z)$ are $\pm i$ (which happens if and only if k is even) and $\pm\rho, \pm\rho^2$ with $\rho = e^{2\pi i/3}$ (which happens if and only if 3 divides k). As a consequence, they deduce that for $k \geq 4$, there is an algebraic α with $|\alpha| = 1$ which is not a $2k$th root of unity so that

$$\zeta(2k+1) = \frac{2}{\alpha^{2k}-1}\big(F_{2k+1}(\alpha) - \alpha^{2k} F_{2k+1}(-1/\alpha)\big).$$

The authors in [65] were led to consider the function

$$G_{2k+1}(z) = \frac{1}{z^{2k}-1}\big(F_{2k+1}(z) - z^{2k} F_{2k+1}(-1/z)\big)$$

since a special value of this function is $\zeta(2k+1)$. They prove the following:

Theorem 6.1 *The set*

$$\big\{G_{2k+1}(z) : \Im(z) > 0, z \in \overline{\mathbb{Q}}, z^{2k} \neq 1\big\}$$

contains at most one algebraic number.

It is clear from these results that Eichler integrals will play a prominent role in our understanding of the special values of $\zeta(2k+1)$.

Apéry's surprising result that $\zeta(3)$ is irrational proved in 1977 seemed to be an isolated result. All attempts to generalize it to $\zeta(5)$ and more generally to $\zeta(2k+1)$ met with limited success. Only recently did Rivoal [173] succeed in showing that infinitely many of $\zeta(2k+1)$ are irrational. For instance, we now know that the dimension of the vector space spanned by the values $\zeta(3), \zeta(5), \ldots, \zeta(2a+1)$ is greater than $c \log a$ for some positive constant c. These results emanate from a subtle study of hypergeometric series along with an irrationality criterion of Nesterenko. The theory of q-series and its interconnection to the theory of modular forms and mock modular forms seem to suggest a new paradigm for research in this context. Indeed, new light is shed on the irrationality questions by studying the q-analogues of the Riemann zeta function. This analog is not unrelated to the functions P, Q, R of Ramanujan. It seems natural to define the q-Riemann zeta functions

$$\zeta_q(s) = \sum_{n=1}^{\infty} \frac{n^{s-1} q^n}{1-q^n}$$

and try to study irrationality and transcendence of these values for specializations of q and s. There are good theoretical reasons for doing so. Indeed, $\zeta_q(2), \zeta_q(4), \zeta_q(6)$ are essentially the same as Ramanujan's P, Q, R respectively. Nesterenko's theorem gives us precise information on transcendence of these values, that is, $\zeta_q(s)$ when s is even, since from the theory of modular forms, $\zeta_q(2k)$ is a polynomial in $\zeta_q(4)$ and $\zeta_q(6)$ for $k \geq 2$. What exactly is the situation with $\zeta_q(2k+1)$?

Recently, Rivoal and others have obtained interesting results in this direction. For instance, Krattenthaler, Rivoal, and Zudilin [101] have shown that the $\mathbb{Q}$-vector space spanned by the values $\zeta_q(1), \zeta_q(3), \ldots, \zeta_q(a)$ has dimension at least $c_1\sqrt{a}$ for some positive constant c_1, provided that q is different from ± 1 and a is an odd integer.

The irrationality of $\zeta_q(1)$ with q algebraic is connected with the study of the q-exponential function and the q-logarithm. The general philosophy of transition from the world of natural numbers to the world of q-numbers is to relate n with $q^n - 1$. Thus, the analog of the q-exponential function is

$$e_q(x) := 1 + \sum_{n=1}^{\infty} \frac{x^n}{(q^n-1)(q^{n-1}-1)\cdots(q-1)},$$

and the q-logarithm is

$$\ell_q(x) := \sum_{n=1}^{\infty} \frac{x^n}{q^n - 1}.$$

It is not difficult to show that

$$e_q(x) = \prod_{n=1}^{\infty} \left(1 + \frac{x}{q^n}\right)$$

and

$$\ell_q(x) = \sum_{n=1}^{\infty} \frac{x}{q^n - x}.$$

One can relate the value of $\zeta_q(1)$ to the special value of a q-logarithm. Even for the q-exponential function, one needs the strength of Nesterenko's theorem to assert that $e_q(1)$ is transcendental for algebraic values of q with $|q| > 1$. It will be interesting to study and develop transcendental number theory to the domain of its q-analogs.

From these results it is clear that this "quantum" world seems to lie deeper than the world of natural numbers. This line of investigation may ultimately shed new light on our understanding of the special values of the Riemann zeta function at odd arguments.

Chapter 7
Arithmetic of the Partition Function

Ramanujan's work on the partition function can be divided into two parts. The first part deals with its arithmetic properties, congruences and identities, now called the Rogers–Ramanujan identities. The second deals with his joint work with Hardy on its asymptotic behaviour, for which the elaborate circle method was developed. Since we discussed the latter method in an earlier chapter, we focus here on the arithmetic side of the partition function. As will be evident, elliptic functions and the modern theory of ℓ-adic representations play a major role in the exposition.

1 Ramanujan's Congruences

Recall that a *partition* of a natural number n is the decomposition of n as a finite number of natural numbers. For example, the five partitions of 4 are given by

$$4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1,$$

so we may think of the summands as being arranged in decreasing order. To the partition $n = \lambda_1 + \lambda_2 + \cdots + \lambda_r$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r$ we may associate a diagram (often called the Young diagram) which has r rows and in the jth row, we place λ_j empty boxes. The number of partitions of n will be denoted by $p(n)$.

At the time of Ramanujan, very little was known about the arithmetic nature of $p(n)$. Even today, though we have a better understanding, there are still many unanswered questions regarding it. For instance, it is still unknown if $p(n)$ is even "half" the time. More precisely, is it true that

$$\lim_{x\to\infty} \frac{\#\{n \leq x : p(n) \text{ even}\}}{x} = \frac{1}{2}?$$

The conjecture is "yes". Looking at values of $p(n)$, one would guess that perhaps $p(n)$ is "random" and no special residue class is favored. Yet, Ramanujan found some surprising congruences:

$$p(5n + 4) \equiv 0 \pmod 5,$$

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_7, © Springer India 2013

$$p(7n+5) \equiv 0 \pmod 7,$$
$$p(11n+6) \equiv 0 \pmod{11}.$$

Surprisingly, one needs the theory of elliptic functions to prove these results. In paper 30 of his Collected Papers, these results were published posthumously in 1921 in Mathematische Zeitschrift, prepared for the journal by G.H. Hardy and based on an unpublished manuscript of Ramanujan.

What is interesting in Ramanujan's paper is that he actually proves more general results and the congruences stated above are corollaries of these general results. However, here we will follow an exposition of Berndt [21] that gives a simplified treatment.

As in Ramanujan [165], we have the Eisenstein series

$$P = 1 - 24\sum_{n=1}^{\infty} \frac{nq^n}{1-q^n},$$
$$Q = 1 + 240\sum_{n=1}^{\infty} \frac{n^3q^n}{1-q^n},$$
$$R = 1 - 504\sum_{n=1}^{\infty} \frac{n^5q^n}{1-q^n}.$$

If D denotes the differential operator qd/dq, then these Eisenstein series satisfy

$$DP = \frac{P^2-Q}{12}, \qquad DQ = \frac{PQ-R}{3}, \qquad DR = \frac{PR-Q^2}{2},$$

as was shown by Ramanujan in his celebrated paper [161] of 1916.

The easiest way of establishing the validity of these differential equations is as follows. If we let $q = e^{2\pi iz}$, then Q and R are modular forms of weight 4 and 6 respectively. P is not a modular form but a quasi-modular form and has the following transformation law. Let $E_2(z) = P(e^{2\pi iz})$ and $G_2(z) = 2\zeta(2)E_2(z)$. Then, we have

$$G_2(z+1) = G_2(z+1), \qquad G_2(-1/z) = z^2G_2(z) - 2\pi iz.$$

Then, it is straightforward to check that

$$DP - (P^2-Q)/12$$

is a modular form of weight 4 which vanishes at infinity. Consequently, it is a cusp form of weight 4. But there are no non-trivial cusp forms of weight 4, and so this is identically zero. The other differential equations are established similarly.

In addition to this, we have the well-known discriminant relation

$$\frac{Q^3-R^2}{1728} = \Delta(q) = q\prod_{n=1}^{\infty}\left(1-q^n\right)^{24}.$$

Following Ramanujan's notation, we will denote by J any q-series with integer coefficients and emphasize that J is not necessarily the same at each occurrence. Since we are interested in congruences, this is a useful device (similar to the big 'O' notation of Landau). We begin with the simplest congruence.

Theorem 1.1

$$p(5m+4) \equiv 0 \pmod 5.$$

Proof Using our notation, we have

$$Q = 1 + 5J, \qquad R = P + 5J,$$

the latter coming from $n^5 \equiv n \pmod 5$, which is a special case of Fermat's little theorem. We have

$$Q^3 - R^2 = Q(1+5J)^2 - (P+5J)^2 = Q - P^2 + 5J.$$

From the differential equations satisfied by P we see that this is

$$= -12DP + 5J = 3DP + 5J.$$

On the other hand,

$$Q^3 - R^2 = 1728q \prod_{n=1}^{\infty} \left(1-q^n\right)^{24} = 3q \prod_{n=1}^{\infty} \frac{(1-q^n)^{25}}{(1-q^n)} + 5J.$$

Using the fact that

$$\left(1-q^n\right)^{25} = 1 - q^{25n} + 5J,$$

we deduce that

$$Q^3 - R^2 = 3q \left(\sum_{n=0}^{\infty} p(n) q^n \right) \prod_{n=1}^{\infty} \left(1-q^{25n}\right) + 5J.$$

Combining this with our earlier formula for $Q^3 - P^2$, we obtain

$$3DP = 3q \left(\sum_{n=0}^{\infty} p(n) q^n \right) \prod_{n=1}^{\infty} \left(1-q^{25n}\right) \pmod 5. \tag{37}$$

To deduce the desired result, we need only make the observation that if J is any integral q-series given by

$$J = \sum_{n=0}^{\infty} a(n) q^n,$$

then

$$DJ = \sum_{n=1}^{\infty} na(n)q^n,$$

so that the nth coefficient is always divisible by n. In (37), we have

$$DP \prod_{n=1}^{\infty} \left(1 - q^{25n}\right)^{-1} \equiv q \sum_{n=0}^{\infty} p(n)q^n \pmod 5,$$

and we compare the coefficients of q^{5m} on both sides of the congruence to deduce the desired result. Indeed, since all the coefficients of q^{5m} of DP are divisible by 5, the same is true of the product on the left-hand side which is an integral q-series. So the same must be true of the right-hand side. But the coefficient of q^{5m} on the right-hand side is precisely $p(5m - 1)$. □

The other congruences are obtained similarly. One needs to use some further identities of the following kind.

Since the q-series of any modular form of weight ≥ 4 and level 1 can be written as a polynomial in Q and R, it is not surprising that we also have

$$Q^2 = 1 + 480 \sum_{n=1}^{\infty} \frac{n^7 q^n}{1 - q^n}, \tag{38}$$

$$QR = 1 - 264 \sum_{n=1}^{\infty} \frac{n^9 q^n}{1 - q^n},$$

and

$$441Q^3 + 250R^2 = 691 + 65520 \sum_{n=1}^{\infty} \frac{n^{11} q^n}{1 - q^n},$$

the right-hand sides being the q-series of Eisenstein series of weight 8, 10, and 12 respectively.

The congruences for 7 and 11 are progressively harder. For the congruence modulo 7, only the identity for Q^2 is needed. For the congruence modulo 11, the other two identities are essential. We proceed with the easier case and follow [21] in our discussion filling in some details.

Theorem 1.2

$$p(7m + 5) \equiv 0 \pmod 7.$$

Proof In our notation, we have

$$R = 1 + 7J,$$

and using Fermat's little theorem (with $p = 7$), from (38) we have

$$Q^2 = P + 7J.$$

Hence, modulo 7, we have

$$\begin{aligned}\left(Q^3 - R^2\right)^2 &\equiv (PQ - 1)^2 \equiv P^2Q^2 - 2PQ + 1\\ &\equiv P^3 - 2PQ + 1 \equiv P\left(P^2 - Q\right) - PQ + R.\end{aligned}$$

Using the differential equations satisfied by P, Q, R, we obtain that

$$\left(Q^3 - R^2\right)^2 \equiv 5P(DP) - 3DQ \pmod 7.$$

This is the same as (modulo 7)

$$6DP^2 - 3DQ.$$

On the other hand,

$$\left(Q^3 - R^2\right)^2 = q^2\prod_{n=1}^{\infty}\left(1 - q^n\right)^{48} = q^2\prod_{n=1}^{\infty}\frac{(1 - q^n)^{49}}{(1 - q^n)}.$$

Reducing this modulo 7, we obtain

$$\left(Q^3 - R^2\right)^2 \equiv q^2\left(\sum_{n=0}^{\infty}p(n)q^n\right)\prod_{n=1}^{\infty}\left(1 - q^{49n}\right) \pmod 7.$$

Combining this with our earlier congruence, we obtain

$$q^2\sum_{n=0}^{\infty}p(n)q^n \equiv \prod_{n=1}^{\infty}\left(1 - q^{49n}\right)^{-1}\left(6q\frac{dP^2}{dq} - 3q\frac{dQ}{dq}\right) \pmod 7.$$

Noting again that for any q-series J with integer coefficients, the nth coefficient of DJ is always divisible by n, we need only compare the coefficients of q^{7n} of both sides of this congruence to deduce the desired result. □

The key point to note here is that all of these congruences are emanating from Bernoulli numbers appearing in the appropriate Eisenstein series. This again connects naturally to the theory of ℓ-adic representations and the Serre–Swinnerton-Dyer theory of congruences.

2 Higher Congruences

Based on very scanty numerical evidence, Ramanujan formulated a more general conjecture [164] modulo powers of primes 5, 7, and 11. In 1930, Chowla [33] and

Gupta (independently) found a counterexample to his conjecture modulo 7^3, so that some correction was needed. The precise version is given by the following:

Theorem 2.1 *Let p be* 5, 7, *or* 11, *and let $\delta_{p,k}$ be the multiplicative inverse of* 24 (mod p^k). *Then*,

(1) $p(5^k m + \delta_{5,k}) \equiv 0 \pmod{5^k}$;
(2) $p(7^k m + \delta_{7,k}) \equiv 0 \pmod{7^{\lfloor k/2 \rfloor + 1}}$;
(3) $p(11^k m + \delta_{11,k}) \equiv 0 \pmod{11^k}$.

In 1938, Watson [199] proved (1) and (2). Atkin [11] proved (3) in 1967. What do these proofs involve? What are the conceptual reasons behind them? Folsom, Kent, and Ono [53] used the theory of ℓ-adic modular forms to give a conceptual explanation of these congruences. We refer the reader to this paper for further details.

3 Dyson's Ranks and Cranks

In 1944, Dyson made an important observation to give combinatorial explanations for these congruences. If $(\lambda_1, \ldots, \lambda_r)$ is a partition λ of n (denoted $\lambda \vdash n$), he defined the *length* of λ, denoted $\ell(\lambda)$, to be r and the *rank* of λ, denoted $r(\lambda)$, to be $\lambda_1 - \ell(\lambda)$. For instance, the partition of 4 given by $(2, 1, 1)$ has length 3 and rank -1. Dyson considered the functions

$$R(n, m, M) = \#\big\{\lambda \vdash n : r(\lambda) \equiv m \pmod{M}\big\}.$$

He conjectured that

$$p(5n + 4) = 5R(5n + 4, m, 5)$$

for every $0 \le m \le 4$ and

$$p(7n + 5) = 7R(7n + 5, m, 7)$$

for every $0 \le m \le 6$. In 1954, using the theory of modular functions, Atkin and Swinnerton-Dyer [13] proved Dyson's conjecture. Unfortunately, Dyson's rank function did not apply to Ramanujan's congruence modulo 11. So he was motivated to conjecture that there is another function, which he called the "crank" function, that would explain combinatorially Ramanujan's congruence (mod 11). In 1988, Andrews and Garvan [8, 56] found Dyson's "crank" function. It is defined as follows. Given a partition λ of n, let $o(\lambda)$ be the number of 1s in λ, and $\mu(\lambda)$ be the number of parts larger than $o(\lambda)$. The *crank* of λ, denoted $c(\lambda)$, is given by λ_1 if $o(\lambda) = 0$ and $\mu(\lambda) - o(\lambda)$ otherwise. It turns out that the crank function gives a combinatorial explanation for Ramanujan's congruences (mod 5) and (mod 7) also. Indeed, define

$$S(n, m, M) = \#\big\{\lambda \vdash n : c(\lambda) \equiv m \pmod{M}\big\}.$$

Theorem 3.1

(1) *For each* $0 \le m \le 4$,

$$p(5n+4) = 5S(5n+4, m, 5).$$

(2) *For each* $0 \le m \le 6$,

$$p(7n+5) = 7S(7n+5, m, 7).$$

(3) *For each* $0 \le m \le 10$,

$$p(11n+6) = 11S(11n+6, m, 11).$$

There has been further work of Garvan, Kim, and Stanton [57] in this context. Ramanujan expected that there are no such simple congruences for primes greater than 11. In searching for a theoretical explanation for Ramanujan's intuition, Ahlgren and Boylan [3] in 2003 proved the following theorem.

Theorem 3.2 *If* ℓ *is a prime and* $0 \le \beta < \ell$ *for which*

$$p(\ell n + \beta) \equiv 0 \pmod{\ell}$$

for every $n \ge 1$, *then*

$$(\ell, \beta) \in \big\{(5,4), (7,5), (11,6)\big\}.$$

A nice exposition is given in [151]. Thus, Ramanujan's congruences for 5, 7, and 11 are the only "simple" congruences for $p(n)$.

Atkin, Klover, Lovejoy, and Ono have done further work extending this result to prime powers. Surprisingly, these extensions involve modular forms of half-integral weight. More precisely, let

$$F(z) = \eta(24z)^{19} = \sum_{n=1}^{\infty} a(n)q^n = q^{19} - 19q^{43} + 152q^{67} + \cdots,$$

$$G(z) = \eta(24z)^{23} = \sum_{n=1}^{\infty} b(n)q^n = q^{23} - 23q^{47} + 230q^{71} + \cdots.$$

If χ_{12} is the quadratic character given by the Kronecker symbol $(12/\cdot)$, then F is a cusp form of weight $19/2$ for $\Gamma_0(576)$ and character χ_{12}. G is a modular form of weight $23/2$ for the same group and character. We then have

$$p\big(5^j n + \delta_{5,j}\big) \equiv 3^{j-1}5^j a(24n+19) \pmod{5^{j+1}}$$

for j odd and

$$p\big(5^j n + \delta_{5,j}\big) \equiv 3^{j-1}5^j b(24n+23) \pmod{5^{j+1}},$$

for j even. These results open a new line of research.

4 Parity Questions

The question of determining the parity of the partition function $p(n)$ was raised by Ramanujan (see [24]), but to date, has not been satisfactorily answered. In 1959, Kolberg [100] proved that $p(n)$ is even infinitely often and odd infinitely often. Since his proof is elementary and short, we give it here. We begin with Euler's pentagonal number theorem:

$$\prod_{n=1}^{\infty}\bigl(1-q^n\bigr)=\sum_{n=-\infty}^{\infty}(-1)^n q^{n(3n+1)/2}.$$

Since

$$\prod_{n=1}^{\infty}\bigl(1-q^n\bigr)^{-1}=\sum_{n=0}^{\infty}p(n)q^n,$$

we deduce the recursion (Euler's identity) for $n\geq 1$,

$$p(n)-p(n-1)-p(n-2)+p(n-5)+\cdots=0,$$

where the general term is given by

$$(-1)^k p\bigl(n-k(3k\pm 1)/2\bigr).$$

If $p(n)$ is even for all $n\geq a$, then with $n=a(3a-1)/2$, the above identity gives

$$p\bigl(a(3a-1)/2\bigr)-p\bigl(a(3a-1)/2-1\bigr)-\cdots\pm p(2a-1)\mp p(0)=0,$$

giving a contradiction. Hence $p(n)$ takes odd values infinitely often.

A similar argument works for even values. Indeed, suppose that $p(n)$ is odd for all $n\geq b$. Taking $n=b(3b+1)/2$ in Euler's identity, we again deduce a contradiction since the left-hand side has an odd number of odd terms.

A mild variation of this argument using Euler's identity leads to the following result. For any given $M>1$ and r arbitrary, there are infinitely many n such that $p(n)\not\equiv r \pmod{M}$. A general conjecture was formulated by M. Newman. Newman's conjecture [144] predicts that for any given r and M, there are infinitely many n such that $p(n)\equiv r \pmod{M}$. In his paper, Kolberg uses Ramanujan's congruences to prove some special cases of this conjecture. Namely, he shows that each of the congruences $p(n)\equiv 0 \pmod{10}$, $p(n)\equiv 5 \pmod{10}$, $p(n)\equiv 0 \pmod{14}$, and $p(n)\equiv 7 \pmod{14}$ has infinitely many solutions.

One can try to derive some quantitative results in this context. The Parkin–Shanks [153]conjecture predicts that $p(n)$ is even "half" the time. As stated earlier, this is still open. There are partial results in this direction. In [148], the authors refine the Kolberg argument and essentially using the pigeonhole principle show that

$$\#\bigl\{n\leq x : p(n) \text{ even}\bigr\}\gg\sqrt{x}.$$

They also count the number of $n \leq x$ for which $p(n)$ is odd. But in this case, they get slightly weaker results of the form $\gg x^{1/2-\epsilon}$ for any $\epsilon > 0$. This was recently improved by Ahlgren [2] and Ahlgren and Boylan [4]. In [3], Ahlgren and Boylan show that Newman's conjecture is true for every prime modulus M, with the possible exception of $M = 3$. In addition, they obtain some quantitative results in this direction.

Subbarao [192] proposed the following generalization. If A, B are integers with $0 \leq B < A$, there are infinitely many n for which $p(An + B)$ is even and infinitely many n for which $p(An + B)$ is odd. Over the decades, several special cases of this conjecture were established. For instance, Hirschhorn and Subbarao [81] proved the conjecture for $A = 16$.

Subbarao's conjecture is practically proved. The "odd" part remains open. More precisely, in a beautiful paper, using the theory of modular forms, Ono [150] has shown that there are infinitely many n such that $p(An + B)$ is even. If for some n_0, we have that $p(An_0 + B)$ is odd, then there are infinitely many n such that $p(An + B)$ is odd. Nicolas, Ruzsa, Sarkozy, and Serre [148] have shown that the number of $n \leq x$ such that $p(An + B)$ is even is $\gg \sqrt{x}$. In the odd case, Ahlgren [2] has shown that the number of $n \leq x$ for which $p(An + B)$ is odd is $\gg \sqrt{x}/\log x$. Recently, some small improvements on these results were obtained by Ahlgren.

There are other arithmetic questions one may ask concerning $p(n)$. We refer the reader to an excellent survey by Adhikari and Mukhopadhay [1] for related questions on this theme. Clearly, there is more work to be done, and what emerges from these recent developments is the ubiquitous modular connection in these problems.

Chapter 8
Some Nonlinear Identities for Divisor Functions

1 A Quadratic Relation Amongst Divisor Functions

Let $\sigma_s(n)$ denote the sum of the sth powers of the divisors of n. Also, set

$$\sigma_s(0) = \frac{1}{2}\zeta(-s)$$

where as usual $\zeta(s)$ denotes the Riemann zeta function. In his fundamental paper "On Certain Arithmetical Functions", Ramanujan showed that there is a quadratic relation involving $\sigma_s(n)$.

We have

$$\sum_{j\le x}\sigma_r(j) = \sum_{j\le x}\sum_{dt=j} d^r.$$

Interchanging the summation, the right-hand side is seen to be

$$\sum_{d\le x} d^r[x/d] = x\sum_{d\le x} d^{r-1} + O\Big(\sum_{d\le x} d^r\Big) = O\big(x^{r+1}\big).$$

On the other hand, we could also write this as

$$\sum_{t\le x}\sum_{d\le x/t} d^r \sim \sum_{t\le x}\frac{1}{r+1}(x/t)^{r+1}.$$

The right-hand side is seen to be

$$\sim \frac{1}{r+1}x^{r+1}\zeta(r+1).$$

The same method will allow us to treat

$$\sum_{j\le x}\sigma_r(j)(x-j)^s$$

and for integers $x = n$, we find that it is

$$\sim \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)n^{r+s+1}$$

provided that $r > 0$ and $s \ge 0$.

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_8, © Springer India 2013

Now let us set

$$\Sigma_{r,s}(n) = \sigma_r(0)\sigma_s(n) + \cdots + \sigma_r(n)\sigma_s(0).$$

Then, using the fact that

$$n^s < \sigma_s(n) < n^s\left(1^{-s} + 2^{-s} + \cdots\right) = n^s\zeta(s)$$

it follows that for $r > 0$ and $s \geq 0$,

$$\liminf \frac{\Sigma_{r,s}(n)}{n^{r+s+1}} \geq \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)$$

and for $r > 0$ and $s > 1$,

$$\limsup \frac{\Sigma_{r,s}(n)}{n^{r+s+1}} \leq \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\zeta(r+1)\zeta(s).$$

Ramanujan proves that whenever r and s are positive odd integers,

$$\Sigma_{r,s}(n) = \frac{\Gamma(r+1)\Gamma(s+1)}{\Gamma(r+s+2)}\frac{\zeta(r+1)\zeta(s+1)}{\zeta(r+s+2)}\sigma_{r+s+1}(n) + \frac{\zeta(1-r)+\zeta(1-s)}{r+s}n\sigma_{r+s-1}(n) + O\left(n^{\frac{2}{3}(r+s+1)}\right).$$

Moreover, he shows that in a finite number of cases, there is no error term. In other words, the O-term above can be removed, and one has an exact identity. The values of (r, s) for which this occurs are given in the table below.

r	s
1	1
1	3
1	5
1	7
1	11
3	3
3	5
3	9
5	7

In particular, the case $r = 1$ and $s = 3$ gives the identity

$$\sigma_1(1)\sigma_3(n) + \sigma_1(3)\sigma_3(n-1) + \sigma_1(5)\sigma_3(n-2) + \cdots + \sigma_1(2n+1)\sigma_3(0) = \frac{1}{240}\sigma_5(2n+1).$$

2 Quadratic Relations Amongst Eisenstein Series

Since the Fourier coefficients of Eisenstein series are given in terms of the functions $\sigma_k(n)$, the relations of the previous section may be interpreted as quadratic relations satisfied by Eisenstein series. Using the fact that the space of modular forms of level

1 and weight k is one-dimensional for $4 \le k \le 10$ and for $k = 14$, one gets the following relations:

$$E_8(z) = E_4(z)^2,$$
$$E_{10}(z) = E_4(z)E_6(z),$$
$$E_{12}(z) - E_8(z)E_4(z) = -\frac{428000}{691}\Delta(z)$$

and

$$E_{12}(z) - E_6(z)^2 = \frac{762048}{691}\Delta(z).$$

All of these identities can be expressed as relations between the functions $\sigma_k(n)$ for various values of k and $\tau(n)$.

Ghate [59] has classified all monomial relations between Eisenstein series.

3 A Formula for the τ-Function

Building on the ideas of the previous section, we can also ask for relations between Eisenstein series and cusp forms. There is, of course, the relation

$$1728\Delta(z) = E_4(z)^3 - E_6(z)^2.$$

We have the Fourier expansion

$$E_k(z) = 1 - \frac{4k}{B_k}\sum_{n\ge1}\sigma_{k-1}(n)q^n.$$

In particular,

$$E_4(z) = 1 + 240\sum_{n\ge1}\sigma_3(n)q^n$$

and

$$E_6(z) = 1 - 504\sum_{n\ge1}\sigma_5(n)q^n.$$

This leads to a nonlinear relation for $\tau(n)$ in terms of generalized divisor functions.

Starting from Lahiri [104], several authors have investigated the possibility of quadratic relations between Eisenstein series and cusp forms. These relations arise by studying the effect of certain differential operators on the space of modular forms. In particular, we will discuss three formulas for the τ-function in terms of divisor functions.

The first of these is due to Niebur [149].

Theorem 3.1 *For $n \ge 1$, we have*

$$\tau(n) = n^4\sigma(n) - 24\sum_{k=1}^{n-1}\big(35k^4 - 52k^3n + 18k^2n^2\big)\sigma(k)\sigma(n-k).$$

Proof Set

$$f(z) = -\log \Delta(z) = -2\pi i z + 24 \sum_{n=1}^{\infty} \frac{\sigma(n) e(nz)}{n}.$$

Now define

$$F(z) = 18\big(f^{(3)}(z)\big)^2 + f'(z) f^{(5)}(z) - 16 f^{(2)}(z) f^{(4)}(z).$$

Then $F(z) = F(z+1)$. Moreover, the first Fourier coefficient is $2^9 \cdot 3 \cdot \pi^6$. Hence, if we show that $F(-1/z) = z^{12} F(z)$, it will follow that

$$F(z) = 2^9 \cdot 3 \cdot \pi^6 \Delta(z).$$

The definition of f implies that

$$f(-1/z) = -\log z^{12} + f(z).$$

Differentiating this repeatedly, we get the identities

$$\begin{aligned}
f'(-1/z) &= -12z + z^2 f'(z), \\
f^{(2)}(-1/z) &= -12z^2 + 2z^3 f'(z) + z^4 f^{(2)}(z), \\
f^{(3)}(-1/z) &= -24z^3 + 6z^4 f'(z) + 6z^5 f^{(2)}(z) + z^6 f^{(3)}(z), \\
f^{(4)}(-1/z) &= -72z^4 + 24z^5 f'(z) + 36z^6 f^{(2)}(z) + 12z^7 f^{(3)}(z) + z^8 f^{(4)}(z), \\
f^{(5)}(-1/z) &= -288z^5 + 120z^6 f'(z) + 240z^7 f^{(2)}(z) + 120z^8 f^{(3)}(z) \\
&\quad + 20z^9 f^{(4)}(z) + z^{10} f^{(5)}(z).
\end{aligned}$$

Using these equations, we find that

$$H(z) = f^{(4)}(z) + f'(z) f^{(3)}(z) - \frac{3}{2}\big(f^{(2)}(z)\big)^2$$

is a cusp form of weight 4 for $SL_2(\mathbb{Z})$ and hence must be identically zero. On the other hand, we have

$$F(-1/z) = z^{12} F(z) - 12z^{11} H'(z) - 48z^{10} H(z)$$

and so

$$F(-1/z) = z^{12} F(z)$$

as required. □

4 Derivatives of Modular Forms

Let f be a modular form of weight k for a group Γ. In particular, we have

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k f(z)$$

for elements

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma. \tag{39}$$

Differentiating this gives

$$(ad-bc)(cz+d)^{-2} f'\left(\frac{az+b}{cz+d}\right) = kc(cz+d)^{k-1} f(z) + (cz+d)^k f'(z)$$

and for $ad-bc=1$, this is

$$f'\left(\frac{az+b}{cz+d}\right) = (cz+d)^{k+1}\big(kcf(z) + (cz+d)f'(z)\big).$$

Thus we see that differentiation in general destroys the modular property.

Define the differential operator

$$D = \frac{1}{2\pi i}\frac{d}{dz}.$$

Maass proposed the operator

$$\delta_k = \frac{1}{2\pi i}\frac{k}{2iy} + D = \frac{1}{2\pi i}\left(\frac{k}{2iy} + \frac{d}{dz}\right)$$

where $z = x + iy$. This operator, when applied to a holomorphic modular form, preserves the modular transformation property but not the holomorphy. Maass [118] proved that if f is a holomorphic modular form of weight k, then $\delta_k f$ is a real-analytic modular form of weight $k+2$. More generally, he considered iterates of this operator, namely

$$\delta_k^{(r)} = \left(\frac{1}{2\pi i}\right)^r \left(\frac{k+2r-2}{2iy} + \frac{d}{dz}\right)\left(\frac{k+2r-4}{2iy} + \frac{d}{dz}\right)\cdots\left(\frac{k}{2iy} + \frac{d}{dz}\right)$$

and showed that it takes holomorphic modular forms of weight k to real-analytic modular forms of weight $k+2r$.

A large number of nonlinear identities can be produced by computing the effect of the Maass operator on Eisenstein series.

In 1956, Rankin [170] proposed a differential operator on pairs of modular forms that preserves holomorphy. Let f be a modular form of weight k, and g a modular form of weight ℓ, both for the same subgroup Γ of $SL_2(\mathbb{Z})$. Now consider the operator

$$[f,g]_n = \sum_{r=0}^{n} (-1)^r \binom{k+n-1}{n-r}\binom{\ell+n-1}{r} D^r f D^{n-r} g.$$

In particular,

$$[f,g]_1 = kf\,Dg - \ell(Df)g.$$

If both f and g are modular for Γ (of weight k and ℓ respectively), then for

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma \tag{40}$$

we have

$$[f,g]_1\left(\frac{az+b}{cz+d}\right)=\frac{1}{2\pi i}(cz+d)^{\ell+k}\big(kf(z)g'(z)-\ell g(z)f'(z)\big) \tag{41}$$

and this is

$$[f,g]_1\left(\frac{az+b}{cz+d}\right)=(cz+d)^{\ell+k+2}[f,g]_1(z).$$

In general, it is a result of Cohen (building on work of Rankin) that $[f,g]_n$ is a modular form of weight $k+\ell+2n$ for the same subgroup Γ. Moreover, if $n\geq 1$, it is a cusp form.

5 Differential Operators and Nonlinear Identities

Many examples of Maass derivatives and Rankin–Cohen brackets have been computed. Such calculations usually result in a nonlinear identity involving Fourier coefficients of Eisenstein series and cusp forms.

Formally, for a series

$$g(z)=\sum_{n=0}^{\infty}a(n)e^{2\pi inz}$$

we have

$$\delta_k^{(r)}g(z)=\sum_{n=0}^{\infty}b(n)e^{2\pi inz}$$

where

$$b(n)=\sum_{j=0}^{r}\binom{r}{j}\frac{\Gamma(k+r)}{\Gamma(k+r-j)}(-4\pi y)^{-j}n^{r-j}.$$

Using this and some general results of Shimura [189, 190], it is shown by Lanphier [109] that

$$\delta_4E_4\cdot\delta_4E_4=\frac{2}{9}\delta_8^{(2)}E_8-\frac{320}{3}\Delta.$$

Taking the projection of both sides to their holomorphic parts, the identity

$$\sum_{m=0}^{n}m\sigma_3(m)(n-m)\sigma_3(n-m)=\frac{2}{9}n^2\sigma_7(n)-\frac{320}{3}\tau(n)$$

is obtained. Rearranging gives the identity

$$\tau(n)=n^2\sigma_7(n)-540\sum_{m=1}^{n-1}m(n-m)\sigma_3(m)\sigma_3(n-m).$$

On the other hand, consider the Rankin–Cohen bracket $[E_4, E_6]_1$, which is a cusp form of weight 12 for the full modular group. Thus, it is a multiple of Δ. Unpacking this gives the relation

$$4E_4DE_6 - 6E_6DE_4 = -3456\Delta. \tag{42}$$

Also, we have the relation

$$E_{10} = E_4E_6.$$

This comes from the fact that the space of modular forms of weight $k \in [4, 10]$ is one-dimensional. Differentiating the above relation gives

$$DE_{10} = E_6DE_4 + E_4DE_6.$$

Inserting this into (42) gives the identity

$$\tau(n) = -\frac{11}{24}n\sigma_9(n) + \frac{35}{24}n\sigma_5(n) + 350\sum_{m=1}^{n-1}(n-m)\sigma_3(m)\sigma_5(n-m).$$

Ramakrishnan and Sahu [167] show how one can obtain all of the results of [109] using Rankin–Cohen brackets in place of the Maass operators and holomorphic projection.

6 Quasi-modular Forms

The Eisenstein series $E_2(z)$ (which is P in Ramanujan's notation) is not quite a modular form. It satisfies the transformation property

$$E_2\left(\frac{az+b}{cz+d}\right) = (cz+d)^2E_2(z) + \frac{12c(cz+d)}{2\pi i}$$

for any

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(\mathbb{Z}).$$

More generally, Kaneko and Zagier [90] introduced the notion of a quasi-modular form to capture this behaviour.

Let k and s denote two non-negative integers. A holomorphic function f on the extended upper half-plane (that is, the upper half-plane together with the cusps) is called a quasi-modular form of weight k and depth s if there are holomorphic functions $Q_0(f), \ldots, Q_s(f)$ with the property that for any

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$$

we have

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^k\sum_{i=0}^{s} Q_i(f)(z)\left(\frac{c}{cz+d}\right)^i.$$

Thus, E_2 is a quasi-modular form for $SL_2(\mathbb{Z})$ of weight 2 and depth 1.

By applying the identity element in the above, we see that a quasi-modular form f satisfies

$$f = Q_0(f).$$

Moreover, quasi-modular forms have Fourier expansions as they are periodic. This can be seen by applying the element

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in SL_2(Z)$$

to a quasi-modular form f giving the relation

$$f(z+1) = Q_0(f)(z) = f(z).$$

Denote by $\tilde{M}_k^{\leq s}$ the space of quasi-modular forms of weight k and depth s for the full modular group. The operator D preserves quasi-modular forms, and in fact

$$D : \tilde{M}_k^{\leq s} \longrightarrow \tilde{M}_{k+2}^{\leq s+1}.$$

Moreover, it is a fact that $s \leq k/2$. At the moment, the only examples of quasi-modular forms that we have are E_2 and derivatives of modular forms. It is a fact that all quasi-modular forms for the full modular group can be obtained in this way. In particular,

$$\tilde{M}_k^{\leq k/2} = \bigoplus_{i=0}^{k/2-1} D^i M_{k-2i} \oplus \mathbb{C} D^{k/2-1} E_2.$$

The functions $Q_i(f)$ in the definition of the quasi-modularity property are themselves quasi-modular forms by a result of Martin and Royer [121], Lemma 19. In fact, $Q_0(f) = f$ and $Q_i(f) \in \tilde{M}_{k-2i}^{\leq s-i}$.

The definition of the Rankin–Cohen bracket can be extended to the space of quasi-modular forms. In particular, Ramanujan's differential equation

$$DE_2 = -\frac{1}{12}\left(E_4 - E_2^2\right)$$

is a consequence of

$$[E_2, \Delta]_1 = \Delta E_4.$$

The relation

$$[E_2, E_2]_4 = -48\Delta$$

gives the Niebur identity

$$\tau(n) = n^4\sigma(n) - 24\sum_{a=1}^{n-1}\left(35a^4 - 52a^3n + 18a^2n^2\right)\sigma(a)\sigma(n-a).$$

Moreover, the relation

$$[E_4, DE_4]_1 = 960\Delta$$

gives the van der Pol identity

$$\tau(n) = n^2\sigma_3(n) + 60\sum_{a=1}^{n-1} a(9a - 5n)\sigma_3(a)\sigma_3(n-a).$$

7 Non-linear Congruences and Their Interpretation

Earlier, we saw that congruences satisfied by the τ-function that related it to various of the σ_k functions could be interpreted in terms of Galois representations. More specifically, the reducibility modulo a prime ℓ of the ℓ-adic representation attached to Δ resulted in a congruence for $\tau(p)$ modulo ℓ for (almost all) primes p.

In the case of non-linear relations, it is also possible to derive congruences by reducing modulo a prime that occurs in the support of the coefficients. For example, the van der Pol identity gives the congruence

$$\tau(n) \equiv n^2\sigma_3(n) \pmod{\ell}$$

for $\ell = 2, 3, 5$ (the primes dividing 60). In particular,

$$\tau(p) \equiv p^2\sigma_3(p) \pmod{\ell}$$

and it is this latter congruence that we seek to "explain".

If we consider the ring of (mod ℓ) Galois representations associated to modular forms, it has an algebra structure under $\otimes$ product. If this algebra is finitely presented, then the set of relations will generate all possible quadratic (and higher-order) relations between the representations and hence also their characters.

Chapter 9
Mock Theta Functions and Mock Modular Forms

1 Historical Introduction

In his last letter to Hardy, written three months before his death in 1920, Ramanujan describes the beginnings of a new theory of what he called "mock theta functions." He gave no precise definition of these objects. Rather, he listed seventeen examples and a qualitative description of their key properties, namely, that these functions have asymptotic expansions at every rational point similar to classical theta functions, but that there was no single theta function whose asymptotic expansion agrees at all rational points with that of the mock theta function.

Since Ramanujan's letter, many mathematicians like G.N. Watson (who had in his possession the famed "lost" notebook), A. Selberg, and G.E. Andrews studied the examples Ramanujan had given and gave rigorous proofs of some of the statements he had made. New excitement emerged in the 1970s when Andrews [7] found Ramanujan's lost notebook among the papers of Watson and attempted to prove some of these conjectures.

Indeed, in 1988, Hickerson (see [78, 79]) proved many of these conjectures. However, no unifying theory emerged from these papers. It was only in 2002, when S. Zwegers completed his doctoral thesis under D. Zagier, that we find in his thesis an intrinsic characterization of mock theta functions. Later, Bringmann and Ono used this theory to resolve several open problems in combinatorics and the theory of q-series.

Before we proceed, it may be useful to direct the reader to the lucid Séminaire Bourbaki survey of Zagier [206] in which he writes "Ramanujan used the word 'theta function' where we would say 'modular form' today, so that 'mock theta functions' meant something like 'mock modular forms'." More precisely, a mock theta function is a mock modular form of weight $1/2$.

All of the seventeen examples of Ramanujan are of the form

$$\sum_{n=0}^{\infty} A_n(q),$$

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_9,

where $A_n(q)$ is in $\mathbb{Q}(q)$, and for all $n \geq 1$, $A_{n+1}(q)/A_n(q) = R(q, q^n)$ for some fixed rational function $R(q, r) \in \mathbb{Q}(q, r)$. Such series are called q-hypergeometric series.

2 Ramanujan's Examples

What exactly was contained in Ramanujan's letter? What were his seventeen examples? Ramanujan divides his examples into three groups, four of "order 3", ten of "order 5" and three of "order" 7, though he gave no definition of what he meant by the word "order". Here are some examples:

$$f(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(1+q)^2 \cdots (1+q^n)^2},$$

$$\phi(q) = \sum_{n=0}^{\infty} \frac{(-q)^{n^2}}{(1+q^2)(1+q^4) \cdots (1+q^{2n})},$$

$$\psi(q) = \sum_{n=1}^{\infty} \frac{(-q)^{n^2}}{(1+q)(1+q^3) \cdots (1+q^{2n-1})}.$$

There is some structural similarity of these functions to those appearing in the Rogers–Ramanujan identities. However, these functions are more complex. Ramanujan states without proof that

$$2\phi(q) - f(q) = f(q) + 4\psi(q) = \frac{1 - 2q + 2q^4 - 2q^9 + \cdots}{(1+q)(1+q^2)(1+q^3)\cdots},$$

where the expression on the right-hand side is easily seen to be a modular form of weight $1/2$ (up to a factor of $q^{-1/24}$). From this example we see there are linear relations among the mock theta functions like $\phi = f + 2\psi$. The space spanned by them contains a subspace of ordinary modular forms, and after multiplying by a suitable power of q, we get modular behaviour. The other mock theta functions in Ramanujan's list have similar properties.

With all of these examples the following features emerge. If $H(q) \in \mathbb{Z}[[q]]$ is one of Ramanujan's mock theta functions, then $H(q)$ acquires some modular transformation property after the following steps: (i) multiply $H(q)$ by a suitable power of q; (ii) change q to $e^{2\pi i\tau}$ and set $h(\tau) = e^{2\pi i\lambda\tau} H(e^{2\pi i\tau})$; (iii) add a simple (but non-holomorphic) correction term so that the corrected function has modular properties.

3 The Work of Zwegers

In 2001, Zwegers (see [206]), in his doctoral thesis, discovered the relation between non-holomorphic modular forms, Lerch sums and indefinite theta series. Zwegers

discovered that mock theta functions are really the "holomorphic parts" of real analytic modular forms of weight 1/2. This allowed one to define higher weight "mock theta functions", now called mock modular forms. Thus, a *mock modular form* is the holomorphic part of a weak Maass form, and there is a simple algebro-analytic description of these objects.

To motivate our discussion, we begin with the Fourier expansion of the Eisenstein series of weight k,

$$E_k(z) = 1 - \frac{4k}{B_k} \sum_{n \geq 1} \sigma_{k-1}(n) q^n,$$

and the observation that the infinite sum can be rewritten as

$$\sum_{d \geq 1} d^{k-1} \sum_{m \geq 1} q^{md} = \sum_{d \geq 1} d^{k-1} \frac{q^d}{1 - q^d}.$$

A more complicated sum that bears some superficial resemblance to this is

$$\sum_{n \in \mathbb{Z}} (-1)^n \frac{q^{n(3n+1)/2}}{(1+q^n)}.$$

This function occurs in Ramanujan's "lost notebook".

Following Zwegers (and the exposition by Zagier), consider the function of two complex variables

$$\mu(u, v) = \mu(u, v, \tau) = \frac{a^{\frac{1}{2}}}{\theta(v)} \sum_{n \in \mathbb{Z}} \frac{(-b)^n q^{n(n+1)/2}}{1 - aq^n},$$

where $q = e^{2\pi i \tau}$, $a = e^{2\pi i u}$ and $b = e^{2\pi i v}$, and $\theta(v)$ is the function

$$\theta(v) = \sqrt{b} q^{1/8} \sum_{m \in \mathbb{Z}} (-1)^m b^m q^{m(m+1)/2},$$

which by the Jacobi identity is

$$\theta(v) = \sqrt{b} q^{1/8} \prod_{n=1}^{\infty} (1 - qn)\big(1 - bq^n\big)\big(1 - b^{-1} q^{n-1}\big).$$

Zwegers defines a function $R(z; \tau)$ with $z, \tau \in \mathbb{C}$ and τ having positive imaginary part, such that the function

$$\mu(u, v, \tau) - R(u - v, \tau)$$

is a Jacobi form. Now using the principle that specializing the elliptic variable in a Jacobi form to a "torsion point" results in a usual modular form (up to a multiple of a power of q), one gets a modular description of mock theta functions.

In particular, the function denoted $f(q)$ in the previous section (which Ramanujan called a mock theta function of order 3) satisfies the identity

$$\prod_{n=1}^{\infty} \big(1 - q^n\big) f(q) = \sum_{n \in \mathbb{Z}} (-1)^n \frac{q^{n(3n+1)/2}}{(1+q^n)},$$

and Zwegers shows that if we set

$$R(\tau) = \sum_{n \equiv 1 \bmod 6} \operatorname{sgn}(n)\beta\big(n^2 y/6\big) q^{-n^2/24}$$

where

$$\beta(x) = \int_x^\infty u^{-\frac{1}{2}} e^{-\pi u},$$

then

$$q^{-1/24} f(q) + R(\tau)$$

transforms like a modular form of weight $1/2$ for the principal congruence subgroup $\Gamma(2)$ of level 2.

4 The Space of Mock Modular Forms

With Zweger's theory in hand, Zagier defines the space $\mathcal{M}_k$ of mock modular forms in the following way. It will contain the space $M_k^!$ of weakly holomorphic modular forms of weight k. These are holomorphic functions on the upper half-plane that transform like modular forms of weight k (under the action of a group Γ) and which may have a pole (of finite order) at each cusp. Then $\mathcal{M}_k$ is an extension that fits into an exact sequence

$$0 \longrightarrow M_k^! \longrightarrow \mathcal{M}_k \longrightarrow M_{2-k} \longrightarrow 0$$

where the first map is inclusion, and the second map to the space M_{2-k} of modular forms of weight $2-k$ (the so-called 'shadow map') is defined as follows. Given a modular form $g \in M_{2-k}$, define the transform

$$g^*(z) = (i/2)^{k-1} \int_{-\overline{z}}^\infty (w+z)^{-k} \overline{g(-\overline{w})}\, dw.$$

This is essentially an Eichler integral. Then the 'shadow' of $h \in \mathcal{M}_k$ is the unique element $g \in M_{2-k}$ for which $h + g^*$ transforms like a modular form of weight k for a certain subgroup of $SL_2(\mathbb{R})$. (The shadow mapping is reminiscent of Serre duality. See also the last section in Chap. 2.)

Neither the existence nor the uniqueness of g is evident from this description. However, there is an alternate description that will make this clear. Denote by M_k the space of real analytic functions f that transform under Γ by

$$f\left(\frac{az+b}{cz+d}\right) = \rho(\gamma)(cz+d)^k f(z)$$

with a character ρ of Γ. Define $\widehat{M_k}$ to be the elements $f \in M_k$ for which $(\Im z)^k \partial F/\partial \overline{z}$ is anti-holomorphic, that is

$$\frac{\partial}{\partial z}\left((\Im z)^k \frac{\partial F}{\partial \overline{z}}\right) = 0.$$

The elements of $\widehat{M_k}$ are called harmonic weak Maass forms. They are non-holomorphic modular forms which are eigenfunctions of the Laplacian with eigenvalue $k(2-k)/4$ and having at most exponential growth at the cusps.

5 Some Applications

With the formulation of the concept of a mock modular form, many arithmetic applications have been developed. Some of these have to do with congruences satisfied by the partition function. Others have to do with central critical values of L-functions. We briefly describe the latter as formulated by Bruinier and Ono [26].

The work of Waldspurger established a relation between the Fourier coefficients of half-integral weight modular forms and central critical values of quadratic twists of the L-functions associated to integral weight modular forms, with the two forms in question being connected by the Shimura correspondence. In the case of Maass forms, work of Katok and Sarnak related Fourier coefficients of Maass forms of weight $1/2$ to line integrals of certain Maass cusp forms. Bruinier and Ono consider the analogues of these results for harmonic weak Maass forms that are not weakly holomorphic modular forms (in other words, whose shadow is nonzero). Such forms have a Fourier expansion of the form

$$f_g(z) = \sum_{n \gg -\infty} c^+(n) q^n + \sum_{n<0} c^-(n) W(2\pi n v) q^n$$

where

$$W(x) = W_k(x) = \Gamma\big(1-k, 2|x|\big)$$

is the incomplete Gamma function. The first sum is referred to as the holomorphic part of f, and the second sum as the non-holomorphic part.

Now consider a normalized eigenform $G(z)$ of weight 2 for $\Gamma_0(p)$ with the property that the sign of the functional equation in its L-function is -1. By Kohnen's theory (which is an explicit form of the Shimura correspondence), there is a newform g of weight $3/2$ for $\Gamma_0(4p)$ in Kohnen's plus space which lifts to G under the Shimura correspondence. Then, Bruinier and Ono prove that there is a weight $1/2$ harmonic weak Maass form f_g (say) on $\Gamma_0(4p)$ with the following property. If we write the Fourier expansion as

$$f_g(z) = \sum_{n \gg -\infty} c_g^+(n) q^n + \sum_{n<0} c_g^-(n) W(2\pi n v) q^n,$$

then for negative fundamental discriminants D for which $(\frac{D}{p}) = 1$, we have the Waldspurger-type formula

$$L(G, \chi_D, 1) = 8\pi^2 \|G\|^2 \|g\|^2 \sqrt{\frac{|D|}{N}} c_g^-(D)^2,$$

where $\|\cdot\|$ denotes the Petersson norm. Moreover, they show that for positive fundamental discriminants D for which $(\frac{D}{p}) = 1$, we have that the derivative $L'(G, \chi_D, 1) = 0$ if and only if $c_g^+(D)$ is algebraic.

Finally, they make interesting conjectures about the number of transcendental Fourier coefficients and the number of non-zero Fourier coefficients.

All of this opens a new world of mathematics. The reader will profit much by reading the excellent survey by Ono [152] to get a glimpse of more wonders. It is thus fitting to close this informal discussion with some prophetic words of Freeman Dyson [45]: "The mock theta functions give us tantalizing hints of a grand synthesis still to be discovered. Somehow it should be possible to build them into a coherent group-theoretical structure, analogous to the structure of modular forms which Hecke built around the old theta functions of Jacobi. This remains a challenge for the future." With these recent advances, Dyson's dream may soon be realized.

Chapter 10
Prime Numbers and Highly Composite Numbers

1 The Divisor Functions

The structural properties of natural numbers and their detailed study forms a significant chapter in analytic number theory. In his memoir on highly composite numbers, Ramanujan initiated an important method to study general arithmetic functions. This method has become a dominant theme in current research. Surprisingly, these studies take us into allied areas of transcendental number theory and an intimate study of the Riemann zeta function.

Ramanujan's memoir of 1915 dealing with highly composite numbers begins with the divisor function $d(n)$ which is the number of divisors of n. It is easy to see that if

$$n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$$

is the unique factorization of n as a product of distinct prime powers, then

$$d(n) = (a_1 + 1)(a_2 + 1) \cdots (a_k + 1).$$

There are many ways one can try to understand the nature of such a function. For example, one can try to understand the *average order* by studying

$$\sum_{n \le x} d(n).$$

This was done by Dirichlet who proved that

$$\sum_{n \le x} d(n) = x \log x + (2\gamma - 1)x + O\left(x^{1/2}\right),$$

where γ is Euler's constant. Thus, on average, $d(n)$ behaves like $\log n$. The celebrated Dirichlet divisor problem is to find the infimum of all θ such that

$$\sum_{n \le x} d(n) = x \log x + (2\gamma - 1)x + O\left(x^{\theta}\right). \tag{43}$$

It is conjectured that this infimum is $1/4$. After Dirichlet, Voronoi proved in 1904 that $\theta \le 1/3$. In 1916, Hardy showed that $\theta \ge 1/4$. After a succession of powerful

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_10, © Springer India 2013

and ingenious methods in the theory of exponential sums, the current record is $\theta \leq 131/416$ due to Martin Huxley in 2003.

One can also study "moments" of the divisor function

$$\sum_{n \leq x} d^k(n)$$

and try to obtain asymptotics for this. This approach was initiated by Ramanujan. For example, he proved that

$$\sum_{n \leq x} d^2(n) = Ax \log^3 x + Bx \log^2 x + Cx \log x + Dx + O\big(x^{3/5+\epsilon}\big),$$

for any $\epsilon > 0$. He noted that the error term can be improved to $O(x^{1/2+\epsilon})$ on the Riemann hypothesis. This should be compared with a result of Ingham [84], who proved in 1927 that

$$\sum_{n \leq x} d(n)d(n+\lambda) \sim \frac{6}{\pi^2}\sigma_{-1}(\lambda)x \log^2 x$$

as x tends to infinity. In other words, $d(n)$ and $d(n+\lambda)$ $(\lambda \neq 0)$ have little "correlation" since this is significantly of lower order by a factor of $\log x$ if $\lambda \neq 0$. More generally, he obtains on the Riemann hypothesis that

$$\sum_{n \leq x} d^k(n) = x P_k(\log x) + O\big(x^{1/2+\epsilon}\big),$$

where $P_k(t)$ is a polynomial of degree $2^k - 1$. In this context, Ramanujan studies the Dirichlet series

$$\sum_{n=1}^{\infty} \frac{d^2(n)}{n^s}$$

and more generally

$$\sum_{n=1}^{\infty} \frac{\sigma_a(n)\sigma_b(n)}{n^s},$$

where $\sigma_a(n) = \sum_{d|n} d^a$ are the generalized divisor functions. This is a precursor to the general construction of a Rankin–Selberg convolution in the theory of automorphic L-functions. Indeed, Ramanujan proves that

$$\sum_{n=1}^{\infty} \frac{\sigma_a(n)\sigma_b(n)}{n^s} = \frac{\zeta(s)\zeta(s-a)\zeta(s-b)\zeta(s-a-b)}{\zeta(2s-a-b)} \tag{44}$$

and, moreover, that

$$\sum_{n=1}^{\infty} \frac{\sigma_a(n)\sigma_b(n)\chi(n)}{n^s} = \frac{L(s,\chi)L(s-a,\chi)L(s-b,\chi)L(s-a-b,\chi)}{L(2s-a-b,\chi^2)}$$

where χ is the non-trivial Dirichlet character (mod 4). Both of these equations are based on a simple power series identity noted by Ramanujan, namely

$$\sum_{n=0}^{\infty}\bigl(1+z+z^2+\cdots+z^n\bigr)\bigl(1+w+w^2+\cdots+w^n\bigr)q^n$$
$$=\frac{1-zwq^2}{(1-q)(1-zq)(1-wq)(1-zwq)}.$$

Estermann [51] considered the more general Dirichlet series

$$\sum_{n=1}^{\infty}\frac{d_k(n)^r}{n^s},$$

where $d_k(n)$ is the number of ways of writing n as a product of k natural numbers. In this context, several cognate sums appear, and they are of two types. The first is of "Rankin–Selberg" type, and the second seems to be of "modular type" at least in a majority of cases. An example of the first type is the sum

$$\sum_{n\le x}d(n)d(n+\lambda),$$

mentioned above, and more generally

$$\sum_{n\le x}\sigma_a(n)\sigma_b(n+\lambda),$$

both of which were studied by Ingham in [84], who derived asymptotic formulas for them. Indeed, their asymptotic behaviour had been conjectured by Ramanujan in his 1915 paper. The second type involves

$$\sum_{j<n}\sigma_a(j)\sigma_b(n-j).$$

These convolutions can be evaluated explicitly in terms of Fourier coefficients of modular forms when a, b are odd and ≥ 3. In fact, Ramanujan showed that the above sum is asymptotic to

$$\frac{\Gamma(a+1)\Gamma(b+1)\zeta(a+1)\zeta(b+1)}{\Gamma(a+b+2)\zeta(a+b+2)}\sigma_{a+b+1}(n)$$

and conjectured that the same must hold for other values of a, b. According to Bruce Berndt's commentary, in 1979, Grosjean [60] disproved this conjecture, but the paper seems lodged in a journal that is not so easily accessible. However, in 1957, Halberstam [66] seems to have also disproved the conjecture using the circle method and estimates for Kloosterman sums. (There is also a later work by Motohashi [127] related to this problem.) Halberstam proves that there are additional terms to the asymptotic formula and when a and b are both odd, these terms disappear.

If at least one of a, b is even, there seem to be no explicit formulas of the type Ramanujan derived. One only has asymptotic formulas. A theoretical explanation of why such formulas are missing would be welcome since there is no apparent modular connection and would perhaps open a new area of research.

2 Ramanujan and the Prime Number Theorem

Ingham [85] noted that Ramanujan's identity (44) implies the non-vanishing of $\zeta(s)$ on $\Re(s) = 1$. It is well known that this is equivalent to the prime number theorem, which asserts that the number of primes $p \leq x$, denoted $\pi(x)$ satisfies the asymptotic law

$$\pi(x) \sim \frac{x}{\log x}$$

as $x \to \infty$. This was proved by Hadamard and de la Vallée Poussin (independently) in 1896. Using an analogue of (44) for L-series, Ingham [85] noticed we also get non-vanishing of Dirichlet L-series on $\Re(s) = 1$, a fact which is equivalent to the prime number theorem for arithmetic progressions. This says that if $\pi(x, q, a)$ is the number of primes $p \leq x$ which are congruent to $a \pmod q$, with $(a, q) = 1$, we have the asymptotic

$$\pi(x, q, a) \sim \frac{x}{\varphi(q) \log x},$$

where φ denotes Euler's function.

As noted by Riemann in his celebrated paper of 1860, the function

$$\operatorname{li} x := \lim_{\epsilon \to 0} \left(\int_0^{1-\epsilon} \frac{dt}{\log t} + \int_{1+\epsilon}^{x} \frac{dt}{\log t} \right)$$

is a better approximation to $\pi(x)$ than $x/\log x$ in the sense that the effective version of the prime number theorem of Hadamard and de la Vallée Poussin can be stated as

$$\pi(x) = \operatorname{li} x + O\left(x e^{-c\sqrt{\log x}}\right)$$

for some constant $c > 0$. Elaborate techniques in the theory of exponential sums have only led to modest improvements of this result. By a 1958 result of Korobov and Richert, we know that

$$\pi(x) = \operatorname{li} x + O\left(x^{-c_1 (\log x)^{2/3}}\right)$$

for some positive constant $c_1 > 0$. There has not been any further improvement. Any result of the form

$$\pi(x) = \operatorname{li} x + O\left(x^{\theta}\right)$$

with $\theta < 1$ is equivalent to the assertion that $\zeta(s) \neq 0$ for $\Re(s) > \theta$. Thus, if we are allowed to take any $\theta > 1/2$, this is equivalent to the Riemann hypothesis. An excellent exposition of the relationship between exponential sums, zero-free regions of $\zeta(s)$ and error terms in the prime number theorem can be found in [198]. In [16], Balasubramanian and Ramachandra show that a suitable extension of (44) immediately leads to the prime number theorem for arithmetic progressions with error terms.

In his first letter to G.H. Hardy sent on 16 January 1913, Ramanujan wrote, "I have found a function which exactly represents the number of primes less than x,

'exactly' in the sense that the difference between the function and the actual number of primes is generally 0 or some finite value even when x becomes infinite. I have got the function in the form of infinite series and have expressed it in two ways:

(1) in terms of Bernoulli numbers ...
(2) as a definite integral ..."

In his second letter, Ramanujan gave a third expression for $\pi(x)$.

These assertions can be given the following form. Using the notation of Hardy [67], let

$$J(x) = \int_0^\infty \frac{(\log x)^t\, dt}{t\zeta(t+1)\Gamma(t+1)},$$

$$G(x) = \frac{4}{\pi}\sum_{k=1}^{\infty} \frac{k}{(2k-1)B_{2k}}\left(\frac{\log x}{2\pi}\right)^{2k-1},$$

$$R(x) = \sum_{k=1}^{\infty} \frac{\mu(k)}{k}\int_c^{x^{1/k}} \frac{dt}{\log t},$$

"where $c = 1.45136380\ldots$ nearly" to quote from Ramanujan's second letter of 27 February 1913.

The last series, $R(x)$, appears for the first time in Riemann's 1860 paper, which Ramanujan could not have read. The series $G(x)$ is related to one discovered by Gram. But the series $J(x)$ was due to Ramanujan. Hardy [67] is of the opinion that Ramanujan discovered all these series for $\pi(x)$ independently.

Ramanujan's assertions imply that if $F(x)$ is any of these three functions,

$$\pi(x) - F(x) = O(1). \tag{45}$$

In his essay on Ramanujan's work on prime numbers, Hardy [67] proceeds to show how (45) is false. Indeed, one can show that

$$J(x) = R(x) + o(1)$$

and

$$G(x) = R(x) + o(1),$$

so that one need only study the validity of

$$\pi(x) = R(x) + O(1).$$

From this would follow

$$\pi(x) = \operatorname{li} x + O\left(x^{1/2+\epsilon}\right)$$

for any $\epsilon > 0$, an assertion already noted to be equivalent to the Riemann hypothesis. However, the result also implies, as indicated by Hardy [67],

$$\pi(x) = \operatorname{li} x - \frac{1}{2}\operatorname{li} x^{1/2} + O\left(x^{1/3+\epsilon}\right)$$

for any $\epsilon > 0$. In particular, this implies

$$\pi(x) - \operatorname{li} x \to -\infty \tag{46}$$

as $x \to \infty$. But in 1914, Littlewood showed that

$$\pi(x) - \operatorname{li} x = \Omega_{\pm}\left(\frac{x^{1/2}\log\log\log x}{\log x}\right).$$

This means that there are positive constants c_1, c_2 such that

$$\pi(x) - \operatorname{li} x > c_1\frac{x^{1/2}\log\log\log x}{\log x}$$

for infinitely many x tending to infinity and

$$\pi(x) - \operatorname{li} x < -c_2\frac{x^{1/2}\log\log\log x}{\log x}$$

for infinitely many x tending to infinity. Thus, (46) is false.

In his letters, Ramanujan also made interesting assertions about what we now call "prime number races." In his first letter, he wrote "The difference between the number of prime numbers of the form $4n-1$ and which are less than x, and those of the form $4n+1$ less than x, is infinite when x becomes infinite." In other words, if we denote by $\pi_a(x)$ the number of primes $p \le x$ with $p \equiv a \pmod 4$, with $a = 1, 3$, then Ramanujan's assertion is that

$$\pi_3(x) - \pi_1(x) \to \infty \tag{47}$$

as x tends to infinity. Now from the prime number theorem for arithmetic progressions we have

$$\pi_1(x) \sim \pi_3(x) \sim \frac{1}{2}\operatorname{li} x$$

as x tends to infinity. It is surprising that Hardy does not discuss this in his book since (47) was shown to be false by Littlewood [112] in 1914.

The theme of "prime number races" is a topic of current research. In the 1960s, Knapowski and Turan [99] wrote a series of papers developing a new branch of number theory called comparative prime number theory to tackle such questions.

However, the study of $\pi_3(x) - \pi_1(x)$ goes back to Tchebycheff [195] in 1853, who made two assertions. Let

$$\Delta(x) = \pi_3(x) - \pi_1(x).$$

(1) Then, there exists an infinite sequence of real numbers $\{x_r\}$ tending to infinity such that

$$\lim_{r\to\infty}\frac{\Delta(x_r)\log x_r}{\sqrt{x_r}} = 1;$$

(2) define the function

$$f(\sigma) = \sum_p (-1)^{(p-1)/2}e^{-\sigma p},$$

where the sum is over prime numbers. Then, as $\sigma \to 0^+$, the function $f(\sigma)$ tends to $-\infty$.

The reader will note that in the summation in (2), primes congruent to 1 (mod 4) are weighted with a + sign and primes congruent to 3 (mod 4) are weighted with a − sign. Ramanujan's assertion can be seen as a variation on assertion (2) of Tchebycheff. Naively, the assertion says that there are "more" primes congruent to 3 (mod 4) than there are primes congruent to 1 (mod 4). As mentioned earlier, Littlewood showed in 1918 that

$$\pi_3(x) - \pi_1(x) = \Omega_{\pm}\left(\frac{x^{1/2}\log\log\log x}{\log x}\right),$$

so that Tchebycheff's assertion (1) is true. In fact, the limit 1 can be replaced by any real number, and the assertion is true. Regarding (2), Hardy and Littlewood [68] show that if χ is the non-trivial character (mod 4) and

$$L(s,\chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

satisfies the analog of the Riemann hypothesis, then there exists an $H > 0$ such that

$$\sum_p (-1)^{(p-1)/2} e^{-\sigma p} < -\frac{H}{\sqrt{\sigma}\log(1/\sigma)}$$

for all sufficiently small values of σ. Thus, (2) follows on the Riemann hypothesis for $L(s,\chi)$.

On the other hand, Landau [105] showed that if $L(s,\chi)$ does not satisfy the Riemann hypothesis, then

$$-\infty = \liminf_{\sigma\to 0^+} f(\sigma) \quad \text{and} \quad \limsup_{\sigma\to 0^+} f(\sigma) = +\infty.$$

Thus, the function in (2) has no limit as $\sigma \to 0^+$ if the Riemann hypothesis for $L(s,\chi)$ is false. Put in another way, (2) implies the Riemann hypothesis for $L(s,\chi)$.

The more general problem of studying

$$\pi(x,q,a_1) - \pi(x,q,a_2)$$

has been explored in detail in a series of papers by Knapowski and Turan [99].

In his letter of February 1913, Ramanujan wrote that

$$\pi(x,q,a_1) - \pi(x,q,a_2) \to \infty$$

for the triples $(q; a_1, a_2) = (4; 3, 1), (6; 5, 1), (8; 3, 1)$ and $(12; 5, 1)$. Of course, all of these assertions are false, but they hint at the fact that Ramanujan gave some thought to such oscillation questions in India, before his journey to England in 1914.

3 Highly Composite Numbers

What is perhaps surprising about the "Twelve Lectures" is the arrangement of topics of Hardy's lectures, as well as the omissions. The first four of the lectures are

devoted to classical analytic number theory. But Ramanujan's paper on highly composite numbers finds no discussion.

A number n is said to by *highly composite* if $d(m) < d(n)$ whenever $m < n$. His paper discussing highly composite numbers is over sixty pages and signals a general method by which to study other arithmetical functions.

To understand Ramanujan's work in this context, we review some history of results concerning the divisor function. It is elementary to show that

$$d(n) = O\left(n^{\epsilon}\right)$$

for any $\epsilon > 0$ (see, for example, [131]). Using the prime number theorem, Wigert [203] proved in 1907 that for any $\epsilon > 0$,

$$d(n) < 2^{(1+\epsilon)\log n/\log\log n} \tag{48}$$

for all n sufficiently large and that

$$d(n) > 2^{(1-\epsilon)\log n/\log\log n} \tag{49}$$

for infinitely many n. This result determines the maximal order of the divisor function. Ramanujan proved (48) and (49) independently without using the prime number theorem. In fact, he proved the stronger assertion that the maximal order of $\log d(n)$ is

$$(\log 2)\,\mathrm{li}(\log n) + O\left((\log n)\exp(-c\sqrt{\log n})\right)$$

for some positive constant $c > 0$. The essential idea is as follows.

Let

$$\theta(x) = \sum_{p \le x} \log p,$$

where the sum is over primes $p \le x$. Suppose that N has k distinct prime factors. Then, it is easy to show that

$$d(N) < \frac{[(\theta(p_k) + \log N)/k]^k}{(\log p_1)\cdots(\log p_k)},$$

where p_j denotes the jth prime number. Now partial summation shows that

$$\pi(x)\log x - \theta(x) = \int_2^x \frac{\pi(t)\,dt}{t}. \tag{50}$$

If we have the prime number theorem, then it is easy to deduce

$$\theta(p_k) \sim p_k \sim k\log k.$$

But one can deduce $\theta(p_k) \sim k\log k$ from (50), and this is essentially Ramanujan's idea (see also Robin [174]).

As noted by Nicolas [146], Ramanujan's method also determines the maximal order of $r(n)$, the number of ways of writing n as a sum of two squares. One finds that $r(n)$ has the same maximal order as the divisor function. Nicolas [147] points out that this result is given incorrectly on p. 270 of the classic textbook by Hardy and Wright [72].

In his paper, Ramanujan introduces the notion of a "superior highly composite number." N is said to be a *superior highly composite number* if for some $\epsilon > 0$, we have

$$\frac{d(m)}{m^\epsilon} \leq \frac{d(N)}{N^\epsilon}$$

for all $m < N$. Since

$$d(m) \leq (m/N)^\epsilon d(N) < d(N)$$

for all $m < N$, we see that a superior highly composite number is also highly composite.

Ramanujan gave a complete classification of highly composite numbers and superior highly composite numbers. Surprisingly, there is an interesting connection to transcendental number theory that we explain in the next section.

4 Relation to the Six Exponential Conjecture

Define

$$\mathcal{E} = \left\{ \frac{\log(1+1/k)}{\log p},\ p \text{ prime}, k \geq 1, k \geq \mathbb{Z} \right\}.$$

If $\epsilon \notin \mathcal{E}$, then the maximum of $d(n)/n^\epsilon$ is attained at only one integer N_ϵ, and we have

$$N_\epsilon = \prod_P p^{\alpha_p}, \quad \alpha_p = \left[1/\left(p^\epsilon - 1\right)\right].$$

By the six exponentials theorem explained below and proved by Siegel, we know that for real λ and three distinct primes p, q, r, the numbers

$$p^\lambda, q^\lambda, r^\lambda$$

cannot all be rational unless λ is an integer. This implies that three elements of $\mathcal{E}$ cannot be equal.

The six exponentials theorem is easy to explain. Suppose that x_1, x_2 are linearly independent over $\mathbb{Q}$ and y_1, y_2, y_3 are linearly independent over $\mathbb{Q}$. Then, at least one of the six exponentials $e^{x_i y_j}$, $1 \leq i \leq 2$ and $1 \leq j \leq 3$, is transcendental. It is conjectured that one can take y_1, y_2 linearly independent over $\mathbb{Q}$ and at least one of the four exponentials $e^{x_i y_j}$ with $1 \leq i, j \leq 2$ is transcendental.

Now suppose that three elements of $\mathcal{E}$ are equal. Then we have

$$\lambda = \frac{\log(1+1/k_p)}{\log p} = \frac{\log(1+1/k_q)}{\log q} = \frac{\log(1+1/k_r)}{\log r}$$

for three distinct primes p, q, r and positive integers k_p, k_q, k_r. Then,

$$p^\lambda = 1 + \frac{1}{k_p}, \quad q^\lambda = 1 + \frac{1}{k_q}, \quad r^\lambda = 1 + \frac{1}{k_r}$$

are all rational. If λ is irrational, this contradicts the six exponentials theorem since we may take $x_1 = 1$, $x_2 = \lambda$ and $y_1 = \log p$, $y_2 = \log q$, $y_3 = \log r$ in the theorem. Therefore, λ is rational. But $p^\lambda = 1 + 1/k_p$ being rational implies that λ must be a positive integer. Now,

$$\lambda = \frac{\log(1 + 1/k_p)}{\log p} \leq \frac{\log 2}{\log p},$$

and hence $p^\lambda \leq 2$, which means $p = 2$ and $\lambda = 1$. The same applies for q and r, a contradiction, since p, q, r are distinct.

The connection to the six exponentials theorem was made by Alaoglu and Erdös [6]. The four exponential conjecture would imply that all elements of $\mathcal{E}$ are distinct. If true, this would mean that $d(n)/n^\epsilon$ attains its maximum at two integers. If false, the maximum would be attained at four integers for some ϵ. As Nicolas [147] points out, this was overlooked by Ramanujan.

5 Counting Highly Composite Numbers

Ramanujan devoted a section of his paper to the study of $Q(x)$, the number of highly composite numbers $\leq x$. Since $d(2n) > d(n)$, we see that between x and $2x$, there is always a highly composite number. Indeed, taking the largest highly composite number v in the interval $[1, x]$, we have $d(v) > d(n)$ for all $n \leq x$, $n \neq v$. Now $2v \in [x, 2x]$ and $d(2v) > d(v)$, so that $Q(x) > \log x$. In his paper, Ramanujan [160] showed that

$$\lim_{x\to\infty} \frac{Q(x)}{\log x} = +\infty.$$

Clearly, the determination of $Q(x)$ was of significant interest to Ramanujan. In their joint paper [70], Hardy and Ramanujan wrote "The problem of determining the number $Q(x)$ of highly composite numbers not exceeding x appears to be one of extreme difficulty. It is still uncertain whether or not the order of $Q(x)$ is greater than that of any power of $\log x$."

Several mathematicians afterwards took up this problem. In 1944, P. Erdös showed that for some $c_1 > 1$, $Q(x) > (\log x)^{c_1}$. The new ingredient was Hoheisel's theorem [82]:

$$\pi\left(x + x^\delta\right) - \pi(x) \geq \frac{x^\delta}{\log x}. \tag{51}$$

Hoheisel's result has since been improved, but the connection to the distribution of primes in short intervals is important. One expects that (51) is true with any $\delta > 0$. The Riemann hypothesis would imply that any $\delta > 1/2$ is permissible.

In 1971, Nicolas [145] proved that there is a constant c_2 such that

$$Q(x) \leq (\log x)^{c_2}.$$

This resolves one of the questions raised by Hardy and Ramanujan. The order of $Q(x)$ is not greater than any power of $\log x$.

Here again, Nicolas [145] used a result from transcendental number theory to prove the theorem. More precisely, he used Feldman's theorem: let

$$\theta = \frac{\log(3/2)}{\log 2}.$$

There exists a κ such that for all integers $u, v \geq 1$, we have

$$|v\theta - u| \geq v^{-\kappa}.$$

The question of determining optimal values of c_1, c_2 remains open. Nicolas remarks that if we assume two conjectures, namely that (51) holds for any $\delta > 0$ and that for all $\eta > 0$, there exists a constant $B(\eta) = B$ such that for all $u, v, w \in \mathbb{Z}$, we have

$$|u \log 5 + v \log 3 + w \log 2| \geq B\big(\big(|u| + 1\big)\big(|v| + 1\big)\big)^{-1-\eta},$$

then writing $Q(x) = (\log x)^{c(x)}$, we have

$$\limsup_{x\to\infty} c(x) = \frac{\log 30}{\log 16} = 1.227\ldots.$$

Unconditionally, we have

$$1.136 \leq c_1 \leq c_2 \leq 1.71.$$

All of these results are based on our knowledge of Diophantine approximations of $(\log(3/2)/\log 2)$.

6 Maximal Order of Divisor Functions and Other Arithmetic Functions

In his paper on highly composite numbers, Ramanujan studied related arithmetical functions like the generalized divisor functions $d_k(n)$ counting the number of ways of writing n as a product of k numbers. He also studied the generalized divisor sum functions $\sigma_s(n)$, the sum of the sth powers of the positive divisors of n. However, these were not published. In fact, sections 53 to 75 of his paper on highly composite numbers were expunged in the final publication since the London Mathematical Society was having financial difficulties. Indeed, in the first edition of Ramanujan's collected papers, the editors wrote, "The paper, as long as it is, is not complete. The London Mathematical Society was in some financial difficulty at the time and Ramanujan suppressed part of what he had written, in order to save expense" (see p. 238 of [145]). Fortunately, the suppressed part has been preserved and has since been published in the first volume of the Ramanujan journal [166]. The sections of the unpublished part of [160] were in a box of G.N. Watson's papers kept at Trinity College Library and accidentally discovered by G.E. Andrews along with the now famous "Lost Notebook" of Ramanujan.

One would wonder at this strange excision of a complete manuscript by the London Mathematical Society. Nicolas and Robin in [166] wrote, "Hardy did not much

like highly composite numbers. In the preface to the "Collected Works" he writes that "The long memoir [160] represents work, perhaps, in a backwater of mathematics," but a few lines later, he does recognize that "it shews very clearly Ramanujan's extraordinary mastery over the algebra of inequalities." One of us can remember Freeman Dyson in Urbana (in 1987), saying that when he was a research student of Hardy, he wanted to do research on highly composite numbers, but Hardy dissuaded him as he thought the subject was not sufficiently interesting or important." This explains to some extent why the topic did not find a place in Hardy's twelve lectures [67]. This is the second occasion in which Hardy used the phrase "backwater of mathematics" referring to some aspect of Ramanujan's work.

In his unpublished sections of [160], Ramanujan extended the notion of a highly composite number to other arithmetical functions such as $r_{2k}(n)$, the number of ways of writing n as a sum of $2k$ squares. He also studied the maximal order of $r_Q(n)$, the number of ways of writing n as $Q(x, y)$ with Q being a positive definite binary quadratic form. As noted above, he considered maximal orders of $d_k(n)$ and $\sigma_s(n)$.

For example, the notion of a superior highly composite number extended to the study of $\sigma(n)$ is a number n such that

$$\frac{\sigma(m)}{m} < \frac{\sigma(n)}{n}$$

for all $m < n$. Such numbers are now called *superabundant numbers*. *Colossally superabundant numbers* are numbers n for which there is a positive $\epsilon > 0$ such that

$$\frac{\sigma(n)}{n^{1+\epsilon}} \geq \frac{\sigma(m)}{m^{1+\epsilon}}$$

for all $m > 1$. It turns out that the set of colossal superabundant numbers is infinite and is a subset of the superabundant numbers [6]. The study of these numbers was initiated by Ramanujan [166] in the suppressed portion of [160].

One particular result here requires special mention. The maximal order of $\sigma_{-1}(n) = \sigma(n)/n$ is $e^{\gamma} \log\log n$, where γ is Euler's constant. From formula (382) of Ramanujan's paper [166] it follows under the Riemann hypothesis that for n_0 sufficiently large,

$$\sigma(n) \leq e^{\gamma} n \log\log n.$$

Robin [174] shows that if this inequality holds for all $n > 5040$, then the Riemann hypothesis follows. Lagarias [103] showed that the Riemann hypothesis is equivalent to

$$\sigma(n) \leq H_n + e^{H_n} \log H_n,$$

where $H_n = \sum_{j \leq n} 1/j$. This particular reformulation finds a place in logic, where it was used to show that the Riemann hypothesis is "Diophantine." In fact, this equivalence of the Riemann hypothesis is intimately connected with the study of highly composite numbers.

7 Maximal Orders of Fourier Coefficients of Cusp Forms

As a consequence of Deligne's work [39], we know that

$$\left|\tau(n)\right| \leq n^{11/2} d(n),$$

and more generally, if $\lambda_f(n)$ is the nth Fourier coefficient of a normalized Hecke eigenform f of weight k, then

$$\left|\lambda_f(n)\right| \leq n^{(k-1)/2} d(n).$$

One would like to determine the maximal order of $\tau(n)$ and generally $\lambda_f(n)$. In other words, are these inequalities sharp? From our earlier discussion about the maximal order of $d(n)$ we see that the above implies

$$\lambda_f(n) = O\left(n^{(k-1)/2} \exp\left(\frac{c \log n}{\log \log n}\right)\right)$$

for some positive constant c. For the full modular group, the reader will recall that normalized Hecke eigenforms have real coefficients, and so, the question of whether these coefficients get very close to the given bound above is the question of its maximal order. R. Murty [128] showed that

$$\lambda_f(n) = \Omega_{\pm}\left(n^{(k-1)/2} \exp\left(\frac{c \log n}{\log \log n}\right)\right)$$

for some $c > 0$. For this purpose, he needed to use analytic properties of symmetric power L-functions attached to the automorphic representation associated to f. With recent progress on the Sato–Tate conjecture, it is likely that new ideas can be injected into the further study of such oscillation results. Indeed, the maximal order of $\tau(n)$ can now be determined explicitly, and one can show that the optimal value of c in the above result is $c = \log 2$.

Chapter 11
Probabilistic Number Theory

1 The Normal Order Method

In their fundamental paper of 1917, Hardy and Ramanujan [69] introduced the concept of "normal order" of an arithmetical function. To be precise, this means the following. Given an arithmetical function $g(n)$ and an increasing function $\phi(n)$, we say that $g(n)$ has *normal order* $\phi(n)$ if for any $\epsilon > 0$, the number of $n \leq x$ such that

$$\left|g(n) - \phi(n)\right| > \epsilon\phi(n) \tag{52}$$

is $o(x)$ as $x \to \infty$. In other words, for almost all natural numbers n (in the sense of natural density), we have

$$(1 - \epsilon)\phi(n) < g(n) < (1 + \epsilon)\phi(n).$$

Some arithmetical functions like the divisor function, for example, may not possess a normal order. But other functions like $\omega(n)$, defined as the number of distinct prime factors of n, do have a normal order. In fact, the main result of [69] is that the normal order of $\omega(n)$ is $\log\log n$.

The paper of Hardy and Ramanujan laid the foundations of a new subject called probabilistic number theory. Since then, the subject expanded into a major discipline at the hands of P. Erdös, M. Kac, J. Kubilius and P.D.T.A. Elliott, to name a few. In fact, by 1972, the subject became so enormous that Elliott wrote a two-volume treatise explaining the developments since the time of the Hardy–Ramanujan paper. In the recent past, there have been further developments especially to the study of Fourier coefficients of modular forms. We summarize these at the end of this chapter.

In their seminal paper, Hardy and Ramanujan make use of Brun's sieve to prove that $\omega(n)$ has normal order $\log\log n$. In 1934, Turan [196] showed how one can derive the Hardy–Ramanujan theorem without Brun's sieve and using what can be viewed as Tchebycheff's inequality. Apparently, this paper of Turan was part of his doctoral thesis written under the direction of L. Féjer, and it seems that Turan was unaware of Tchebycheff's inequality or probability theory. As the proof is quite simple, we reproduce it here.

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_11, © Springer India 2013

It is fairly straightforward to show that

$$\sum_{n\le x}\omega(n) = x\log\log x + O(x)$$

and

$$\sum_{n\le x}\omega^2(n) = x(\log\log x)^2 + O(x\log\log x).$$

Putting these two together, we get

$$\sum_{n\le x}\big(\omega(n) - \log\log n\big)^2 = O(x\log\log x).$$

Thus, for any function θ with $\theta(n)\to\infty$ as $n\to\infty$, we have the number of $n\le x$ for which $|\omega(n)-\log\log n| > \theta(n)(\log\log n)^{1/2}$ is $O(x/\theta(x))$, which is $o(x)$.

One can actually sharpen this and create a sieve method. This was done in [113] and [115] and expanded to a combinatorial setting in [114]. The method therefore has immense potential and applications beyond number theory.

2 The Erdös–Kac Theorem

It is interesting that Hardy and Ramanujan chose the word "normal" to describe (52) since, as it transpired later on, there is an intimate connection to the normal distribution in probability theory. To elaborate, let $\theta(n)$ be any function tending to infinity as n tends to infinity. Hardy and Ramanujan proved that the number of $n\le x$ such that

$$\big|\omega(n) - \log\log n\big| > \theta(n)(\log\log n)^{1/2} \tag{53}$$

is $o(x)$

In 1940, Erdös and Kac [49] proved a sharper result. Namely, they showed that if

$$N(x) = \#\left\{n\le x : \alpha \le \frac{\omega(n)-\log\log n}{\sqrt{\log\log n}} \le \beta\right\},$$

then

$$\lim_{x\to\infty}\frac{N(x)}{x} = \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt.$$

Thus, the distribution of values of $\omega(n)$ has an intimate connection to the normal distribution.

Erdös and Kac [49] proved a more general theorem. To explain this, we recall some definitions. An arithmetical function f is called *additive* if $f(mn) = f(m)+f(n)$ for all coprime m,n. An additive function is called *strongly additive*

if $f(p^a) = f(p)$ for p prime and all natural numbers a. If f is a strongly additive function satisfying $|f(p)| \leq 1$ for all primes p, let

$$A_n = \sum_{p \leq n} \frac{f(p)}{p}$$

and

$$B_n = \left(\sum_{p \leq n} \frac{f^2(p)}{p} \right)^{1/2}.$$

Let

$$N_f(x) = \#\left\{ n \leq x : \alpha \leq \frac{f(n) - A_n}{B_n} \leq \beta \right\}.$$

Then,

$$\lim_{x \to \infty} \frac{N_f(x)}{x} = \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-t^2/2}\, dt.$$

Later, Kubilius and Shapiro (independently) (see Theorem 12.2 of [46]) generalized the Erdös–Kac theorem by eliminating the condition that $|f(p)| \leq 1$ and replacing it with the condition that for each $\epsilon > 0$,

$$\lim_{n \to \infty} \frac{1}{B_n^2} \sum_{p \leq n; |f(p)| > \epsilon B_n} \frac{f(p)^2}{p} = 0.$$

It turns out that the same conclusion now holds.

3 The Hardy–Ramanujan-Type Theorem for the τ-Function

The theorem of Deligne connecting the theory of ℓ-adic representations to Fourier coefficients of Hecke eigenforms opened the door for obtaining many new results regarding the arithmetical nature of these coefficients. In particular, one can now study the number of prime factors of $\tau(p)$ and other arithmetical functions arising from modular forms. Unfortunately, to establish these results, some form of a quasi-Riemann hypothesis needs to be assumed for the non-abelian Artin L-series associated to the ℓ-adic representation. We explain why.

More precisely, this theorem says that if $G_{\mathbb{Q}}$ is the absolute Galois group of $\overline{\mathbb{Q}}$ over $\mathbb{Q}$, then there exists an ℓ-adic representation

$$\rho_\ell : G_{\mathbb{Q}} \to GL_2(\mathbb{Z}_\ell),$$

where $\mathbb{Z}_\ell$ denotes the ring of ℓ-adic integers such that if σ_p denotes the Frobenius automorphism, then

$$\operatorname{tr}\big(\rho_\ell(\sigma_p)\big) = \tau(p), \quad \det \rho_\ell(\sigma_p) = p^{11}, \quad p \neq \ell.$$

In particular, if we want to study the distribution of values of $\tau(p)$ modulo ℓ^n, then this means that we can study the finite Galois extension K_{ℓ^n} and the representation,

$$\rho_{\ell,n} : \mathrm{Gal}(K_{\ell^n}/\mathbb{Q}) \to GL_2\big(\mathbb{Z}/\ell^n\mathbb{Z}\big).$$

Ribet has analyzed the image of this representation. It consists of matrices whose determinant is an 11th power (mod ℓ^n).

Using this representation, the study of $\tau(p) \pmod{\ell}$ is reduced to the study of the distribution of the Frobenius automorphism in $\mathrm{Gal}(K_\ell/\mathbb{Q})$. But this distribution is given by the Chebotarev density theorem. Assuming the GRH allows us to use an effective version of this theorem to control error terms that arise in the analysis. It turns out that the full strength of the GRH is not essential in such theorems. Rather, a quasi-GRH, which assumes a zero-free region for the zeta functions in some fixed half-plane to the left of $\Re(s) = 1$ is sufficient for our purpose.

In [132], assuming a quasi-GRH, we showed that

$$\sum_{p\le x}{}' \big(\omega\big(\tau(p)\big) - \log\log p\big)^2 = O\big(\pi(x)\log\log x\big),$$

where the dash on the summation indicates we sum over primes p for which $\tau(p) \neq 0$. Under the same hypothesis, we proved that

$$\sum_{n\le x}{}' \left(\omega\big(\tau(n)\big) - \frac{1}{2}(\log\log n)^2\right)^2 \ll x(\log\log x)^3(\log\log\log\log x).$$

In particular, for any $\epsilon > 0$, we have

$$\big|\omega\big(\tau(p)\big) - \log\log p\big| < (\log\log p)^{1/2+\epsilon}$$

for all but $o(x/\log x)$ primes $p \le x$.

There is nothing particular about $\tau(p)$ or $\tau(n)$. One can establish similar results for Fourier coefficients of normalized Hecke eigenforms of arbitrary weight and level. These results can be viewed as the modular analogues of the classical theorem of Hardy and Ramanujan.

In [133], we extended these results and showed that

$$\#\left\{p \le x, \tau(p) \neq 0 : \alpha \le \frac{\omega(\tau(p)) - \log\log p}{(\log\log p)^{1/2}} \le \beta\right\} \sim \pi(x)\Phi(\alpha,\beta),$$

where

$$\Phi(\alpha,\beta) = \frac{1}{\sqrt{2\pi}}\int_\alpha^\beta e^{-t^2/2}\,dt.$$

Also,

$$\#\left\{n \le x : \alpha \le \frac{\omega(\tau(n)) - \frac{1}{2}(\log\log n)^2}{(\log\log n)^{3/2}/\sqrt{3}} \le \beta\right\} \sim x\Phi(\alpha,\beta),$$

with both theorems using a quasi-GRH.

Of course, analogous results hold for Fourier coefficients of any Hecke eigenform with rational integer coefficients.

4 Non-abelian Generalizations of the Hardy–Ramanujan Theorem

The methods used to derive modular analogues of the Hardy–Ramanujan theorem extend to the study of other arithmetical functions which are not necessarily Fourier coefficients of modular forms or functions related to an ℓ-adic representation. Indeed, if we have a multiplicative function f such that the set of primes p for which $\ell | f(p)$ can be described by Chebotarev conditions, then one can apply similar methods to derive analogous results.

A good example is the case where $f(n) = \phi(n)$, Euler's ϕ-function. In this setting, one can study $\omega(f(p))$ which is the same as the study of $\omega(p-1)$, and one can also study $\omega(\phi(n))$. This was first done by the authors in [132]. Independently and later, Erdös and Pomerance [50] proved that the number of $n \leq x$ for which

$$\alpha \leq \frac{\omega(\phi(n)) - \frac{1}{2}(\log\log n)^2}{(\log\log n)^{3/2}/\sqrt{3}} \leq \beta$$

is asymptotically $x\Phi(\alpha, \beta)$. Here the use of the quasi-GRH is not needed, and one can use the Bombieri–Vinogradov theorem to make the result unconditional.

For the method to work, it is not even essential to have a multiplicative function. To study the divisibility of any sequence of numbers, it is possible to apply these techniques as long as one can relate it again to the Chebotarev density theorem. We give one such example of this type of theorem.

Let a be a natural number which is not equal to ± 1 or a perfect square. A classical conjecture of Artin predicts that a is a primitive root (mod p) for infinitely many primes p. This problem is still open though some spectacular progress has been made in the recent past. In this context, it is natural to study $f_a(p)$, the order of a mod p. Let $f_a(n)$ be the order of a (mod n) with n coprime to a. Under the quasi-GRH, M. Ram Murty and his student F. Saidak [136] showed that the number of primes $p \leq x$ for which

$$\alpha \leq \frac{\omega(f_a(p)) - \log\log p}{\sqrt{\log\log p}} \leq \beta$$

is asymptotic to $\pi(x)\Phi(\alpha, \beta)$. Similarly, under the same hypothesis, the number of $n \leq x$ for which

$$\alpha \leq \frac{\omega(f_a(n)) - \frac{1}{2}(\log\log n)^2}{(\log\log n)^{3/2}/\sqrt{3}} \leq \beta$$

is asymptotic to $x\Phi(\alpha, \beta)$ as x tends to infinity.

Clearly, the normal order method is a powerful method to analyze the arithmetic structure of complicated arithmetical functions. Its potential has not yet been fully exhausted.

Chapter 12
The Sato–Tate Conjecture for the Ramanujan τ-Function

1 Introduction

It was Serre [185] who first conjectured that the Sato–Tate distribution holds for the Ramanujan τ-function. More precisely, he conjectured that the numbers

$$\frac{\tau(p)}{2p^{11/2}},$$

as one ranges over primes p, are equidistributed in the interval $[-1, 1]$ with respect to the Sato–Tate measure

$$\frac{2}{\pi}\sqrt{1-x^2}\,dx.$$

In fact, Serre expected that an analogous result should hold for the Fourier coefficients of any normalized Hecke eigenform not of CM type. This was recently proved by Barnet-Lamb, Geraghty, Harris and Taylor [18]. In particular, this applies to Hecke eigenforms of weight 2 not of CM type. By the celebrated modularity conjecture, we know that elliptic curves correspond to Hecke eigenforms of weight 2 with integer Fourier coefficients. Consequently, the "angles of Frobenius" for elliptic curves not of CM type are distributed with respect to the Sato–Tate measure. This was originally proved for elliptic curves with non-integral j-invariant in [35], but the condition was removed in the most recent work [18].

In this chapter, we give an outline of the proof of this theorem. Since the prerequisites are formidable, we aim only for conceptual clarity and arrange the main ideas for the appreciation and understanding of the reader. We also indicate how one can generalize the theorem so as to introduce additional constraints on the primes, such as primes in a fixed arithmetic progression. In particular, for any interval I in $[-1, 1]$, we will show that for a given natural number q and a coprime to q, the density of primes $p \equiv a \pmod{q}$ with $\tau(p)/2p^{11/2} \in I$ is

$$\frac{2}{\pi\varphi(q)}\int_I \sqrt{1-x^2}\,dx.$$

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2_12, © Springer India 2013

We begin with a historical description of the original conjecture by Sato and Tate in the context of elliptic curves.

Consider the elliptic curve E defined by the equation

$$y^2 = x^3 + ax + b, \quad a, b \in \mathbb{Z}.$$

Let $\Delta = -16(4a^3 + 27b^2)$. As discussed in greater generality in Chap. 3, for each prime p with $(p, \Delta) = 1$, we may consider the congruence

$$y^2 \equiv x^3 + ax + b \pmod{p}$$

and count the number N_p of solutions (x, y). This quantity was first studied by Emil Artin [10] in his 1924 doctoral thesis in which he conjectured that

$$|N_p - p| \leq 2\sqrt{p} \tag{54}$$

for all such primes. In many ways, his study was motivated by the classical Riemann hypothesis. To understand the nature of zeta functions in general, Artin defined the analogue of the Dedekind zeta function in the setting of a function field over a finite field. In the case of a quadratic extension of $\mathbb{F}_p(x)$ defined by

$$y^2 = x^3 + ax + b,$$

the analogue of the Riemann hypothesis for the function field zeta function turns out to be equivalent to (54). In his thesis, Artin verified his conjecture for many small primes p but could not prove it. In February 1933, Hasse [74] proved the conjecture using techniques from algebraic geometry. One could say that understanding this function field analogue of the Riemann hypothesis was an important step in the annals of mathematics. The reader is referred to the historical document [54] for further discussions of this.

Artin's thesis was seminal in many ways. First, it opened up the study of algebraic geometry over finite fields and connected it to the study of exponential sums that occur in classical analytic number theory. Second, it inspired Weil [201] to formulate in 1949 general conjectures that led Grothendieck [63] to chart out a visionary program in algebraic geometry ultimately leading to a resolution of the Weil conjectures in the fundamental work of Deligne [39] in 1974.

Around 1960, Mikio Sato and John Tate [193] (independently) asked about the distribution of the numbers

$$\frac{N_p - p}{\sqrt{p}}$$

in the interval $[-2, 2]$ as p tends to infinity. For example, is it reasonable to expect that these numbers are uniformly distributed in the interval? In other words, is it true that for any interval $[a, b] \subseteq [-2, 2]$, we have

$$\lim_{x \to \infty} \frac{\#\{p \leq x : (N_p - p)/\sqrt{p} \in [a, b]\}}{\#\{p \leq x\}} = b - a?$$

This question is the genesis of the Sato–Tate conjecture. Numerical evidence seemed to suggest otherwise. More precisely, Sato and Tate were led to predict that for a "generic" elliptic curve E, the following is true. If we write

$$(N_p - p)/\sqrt{p} = 2\cos\theta_p, \quad 0 \le \theta_p \le \pi,$$

and $[\alpha, \beta] \subseteq [0, \pi]$, then, their conjecture says

$$\lim_{x\to\infty} \frac{\#\{p \le x : \theta_p \in [\alpha, \beta]\}}{\#\{p \le x\}} = \frac{2}{\pi}\int_\alpha^\beta \sin^2\theta\, d\theta = \frac{\beta-\alpha}{\pi} - \frac{1}{2\pi}(\sin 2\beta - \sin 2\alpha).$$

The conjecture of Tate [193] on algebraic cycles also provided conceptual evidence for this. Here, "generic" means that the elliptic curve should be without complex multiplication (see [107] for details). In other words, E is said to be without complex multiplication (or without CM) if $\operatorname{End} E = \mathbb{Z}$. If the elliptic curve has complex multiplication, then the (essentially) uniform distribution law for the angles was worked out by Deuring [41] building on earlier work of Hecke [75, 76].

One can formulate a more general conjecture. Let E be an elliptic curve defined over a number field K. For each place v of K where E has good reduction, we may consider the group of points of E mod v. Its cardinality (including the identity element) can be written as

$$Nv + 1 - a_v,$$

where Nv denotes the norm of v, and a_v is an integer satisfying Hasse's inequality $|a_v| \le 2(Nv)^{1/2}$. As before, one can therefore write

$$a_v = 2N(v)^{1/2}\cos\theta_v,$$

where $\theta_v(E) := \theta_v$ satisfies $0 \le \theta_v \le \pi$. The Sato–Tate conjecture now is a statement about how the angles θ_v are distributed in the interval $[0, \pi]$ as v varies. When E has complex multiplication (CM), the distribution law is known and again follows from the classical work of Deuring on Hecke L-series (see [139] for details). In the non-CM case, one expects that the angles are uniformly distributed with respect to the measure

$$\mu_{ST} := \frac{2}{\pi}\big(\sin^2\theta\big)\, d\theta.$$

On 18 March 2006, Taylor [194] (see also [35]) published a proof of this conjecture when E has at least one prime of multiplicative reduction. He was building on his earlier work with Clozel, Harris and Shepherd-Barron (see Carayol's Séminaire Bourbaki article [30]).

Inspired by the Sato–Tate conjecture for elliptic curves, Serre [185] extended the conjecture for Fourier coefficients of normalized Hecke eigenforms. We review his formalism below and in the next sections give an axiomatic exposition that allows for an application of results by Barnet-Lamb, Geraghty, Harris and Taylor [18] as well as results pertaining to Hilbert modular forms by Barnet-Lamb, Gee and Geraghty [17]. These new developments all prove special cases of the following hypothesis.

Potential automorphy hypothesis $H_K(S, T, \pi)$ Let K be an algebraic number field, and π a cuspidal automorphic representation of $GL_2(\mathcal{A}_K)$ not of CM type. Let S be a finite set of primes of K, and T a finite set of odd positive integers. Then, there exists a finite Galois extension $K_{S,T}$ over K unramified at S with $\mathrm{Sym}^m(\pi)$ automorphic over $K_{S,T}$ for all $m \in T$.

Our first result is a hybrid Chebotarev–Sato–Tate theorem.

Theorem 1.1 *Let K be an algebraic number field. Let π be a cuspidal automorphic representation of $GL_2(\mathbb{A}_K)$, which is not of CM type, M/K a finite solvable Galois extension with $G = \mathrm{Gal}(M/K)$, and C a conjugacy class of G. For every odd number m, assume hypothesis $H_K(S, T, \pi)$ where S is the set of primes of K that ramify in M and T is either $\{m\}$ or $\{1, m\}$. Then the density of prime ideals $\mathfrak{p}$ for which the Artin symbol $\sigma_\mathfrak{p} \in C$ and the angle $\theta_\mathfrak{p} \in [\alpha, \beta]$ with $0 \le \alpha \le \beta \le \pi$ is*

$$\frac{2|C|}{\pi|G|} \int_\alpha^\beta \sin^2\theta \, d\theta.$$

By the work of [18], the potential automorphy hypothesis is now a theorem in the case of classical normalized Hecke eigenforms not of CM type. In particular, we have the following corollary:

Corollary 1.2 *Let f be a normalized Hecke eigenform of weight k and arbitrary level, which is not of CM type. Write the pth Fourier coefficient of f as $a_p(f) = 2p^{(k-1)/2}\cos\theta_p(f)$. Let q be a natural number, and a an integer with $(a, q) = 1$. For $0 \le \alpha \le \beta \le \pi$, the density of primes p for which $\theta_p(f) \in [\alpha, \beta]$ and $p \equiv a \pmod q$ is*

$$\frac{2}{\pi\varphi(q)} \int_\alpha^\beta \sin^2\theta \, d\theta.$$

It is evident that by similar arguments, one can handle the joint distribution of angles of any finite set of Hecke eigenforms, provided that there is at most one eigenform in the set without CM.

One can also study the joint distribution of angles of a finite collection of pairwise non-isogenous elliptic curves. This looks like a difficult question, and Harris has recently announced some progress in this direction.

Though our treatment is informal, the background needed for a total understanding is quite formidable spanning representation theory, arithmetic algebraic geometry, analytic and algebraic number theory. Still, we hope that the presentation given here will enable the non-expert to see how the proof is put together and appreciate the interplay of ideas.

Recently, "friendly" exposés of the Sato–Tate conjecture have appeared in various places. The papers by Mazur [123, 124] are a good place to begin for the totally uninitiated reader. Here, our goal is more mathematical. We hope to give (without proofs) the main mathematical ingredients that enter into such equidistribution

questions so that the reader may have a conceptual understanding of the results. Because of the celebrated modularity conjecture (now proved by Wiles [204] and others [25]), one can deduce the original conjecture of Sato and Tate, at least in the case that E is defined over the rational number field. In the general number field case, it is still open whether the a_vs can be viewed as coming from automorphic representations as predicted by Langlands. Thus, one can view the Sato–Tate conjecture for elliptic curves as a special case of a more general statement concerning distribution of eigenvalues of Hecke operators or in the general case, of Satake parameters. (See [36] for more details.) One can formulate a function field analogue of the Sato–Tate conjecture, and this has been proved in various contexts. Let K be a rational function field in one variable over a finite field $\mathbb{F}$, and let E denote an elliptic curve over K with nonconstant j-invariant. Let Y denote the projective line over $\mathbb{F}$ and consider the Néron model $\mathcal{E} \longrightarrow Y$. This is a smooth group scheme whose general fibre is E and outside of a finite set S of points $y \in Y$, the fibre $\mathcal{E}_y$ at y is an elliptic curve (the "reduction" of E modulo the residue field corresponding to y). Thus, as an elliptic curve over a finite field, the zeta function of $\mathcal{E}_y$ is of the form

$$\frac{(1-\alpha_y T)(1-\overline{\alpha_y} T)}{(1-T)(1-q^{\deg y} T)}.$$

Here $\alpha_y = q^{(\deg y)/2} e^{i\theta(y)}$ where $0 \le \theta(y) \le \pi$. Let $\mathbb{F}_n$ denote the unique extension of $\mathbb{F}$ of degree n. Then, the Sato–Tate conjecture in this context says that as $n \longrightarrow \infty$, we have

$$\#\big\{y \in Y(\mathbb{F}_n) : \alpha \le \theta_y \le \beta\big\} \sim \left(\int_\alpha^\beta \frac{2}{\pi} \sin^2\theta \, d\theta\right) \big|Y(\mathbb{F}_n)\big|.$$

This was proved by Yoshida in [205] and in a different way by K. Murty in [139]. Very general theorems of Sato–Tate type (in which the base Y is replaced by an arbitrary variety and $\mathcal{E}$ by families of ℓ-adic sheaves) are proved in Deligne [40], Sect. 3.5.

2 Weyl's Criterion

We will begin with a general discussion of the classical setting for uniform distribution. A sequence of real numbers $\{x_n\}$ is called *uniformly distributed* (modulo 1) if for any pair of real numbers α, β with $0 \le \alpha < \beta < 1$, we have

$$\#\big\{n \le N : x_n \in (\alpha, \beta)\big\} \sim (\beta - \alpha)N$$

as N tends to infinity.

Theorem 2.1 (Weyl's Criterion) *The sequence* $\{x_n\}$ *is uniformly distributed* mod 1 *if and only if for all* $m \geq 1$,

$$\sum_{n \leq N} e^{2\pi i m x_n} = o(N)$$

as N *tends to infinity.*

Proof (Sketch) First suppose that the sequence is uniformly distributed. We will show the condition is necessary. Let us observe that any continuous function f can be approximated by a linear combination of step functions so that for any given $\epsilon > 0$, we have

$$\sup_{x \in [0,1]} \left| f(x) - \sum_i c_i \chi_{I_i}(x) \right| \leq \epsilon,$$

where χ_I denotes the characteristic function of an interval I. Then,

$$\sum_{n \leq N} f(x_n) = \sum_i c_i \left(\sum_{n \leq N} \chi_{I_i}(x_n) \right) + O(\epsilon N).$$

By hypothesis,

$$\sum_{n \leq N} \chi_{I_i}(x_n) = \mu(I_i) N + o(N),$$

where $\mu(I)$ denotes the measure of an interval I. Now the sum

$$\sum_i c_i \mu(I_i)$$

is a Riemann sum, and as our ϵ gets smaller, the sum converges to the integral

$$\int_0^1 f(x)\, dx.$$

Thus,

$$\frac{1}{N} \sum_{n \leq N} f(x_n) \to \int_0^1 f(x)\, dx.$$

In particular, we can apply this to $\cos mx$ and $\sin mx$ to deduce the required result.

For the converse, we approximate $\chi_I(x)$ by trigonometric polynomials (which can be done by the Stone–Weierstrass theorem). In fact, one can be more precise. For any positive integer K, there are trigonometric polynomials $m(x)$ and $M(x)$ of degree $\leq K$ such that

$$m(x) \leq \chi_I(x) \leq M(x)$$

with

$$m(x) = \sum_{|m| \le K} a_m e^{2\pi i m x}, \qquad M(x) = \sum_{|m| \le K} b_m e^{2\pi i m x}$$

with

$$a_0 = b_0 = \mu(I) + O(1/K).$$

Therefore,

$$\#\{n \le N : x_n \in I\} = \sum_{n \le N} \chi_I(x_n) = \mu(I)N + o(N),$$

as required. □

Theorem 2.1 says that to establish uniform distribution of the angles θ_v, we need to study the exponential sums

$$\sum_{N(v) \le x} e^{2\pi i m \theta_v}.$$

Weyl's theorem is the classical theorem regarding equidistribution for the group $\mathbb{R}/\mathbb{Z}$. Serre [185] extended the Weyl criterion for arbitrary compact groups and formulated an L-series formalism related to it. We review this in Sect. 4. In the next section, we review a classical Tauberian theorem that will be essential to our discussion.

3 Wiener–Ikehara Tauberian Theorem

Theorem 3.1 *Let $f(s) = \sum_{n=1}^{\infty} a_n/n^s$ with $a_n \ge 0$ and $g(s) = \sum_{n=1}^{\infty} b_n/n^s$ be two Dirichlet series with $|b_n| \le a_n$ for all n. Assume that $f(s)$ and $g(s)$ extend analytically to $\Re(s) \ge 1$ except possibly at $s = 1$ where they have a simple pole with residues R and r (which may be zero) respectively. Then*

$$\sum_{n \le x} b_n \sim rx$$

as x tends to infinity.

The classical application of this theorem is the deduction of the prime number theorem. Let

$$f(s) = -\frac{\zeta'}{\zeta}(s) = \sum_{n=1}^{\infty} \frac{\Lambda(n)}{n^s}$$

where $\Lambda(n) = \log p$ when $n = p^a$ for some prime p and zero otherwise. Taking $g(s) = f(s)$ in the above theorem allows us to deduce the prime number theorem

$$\sum_{n \le x} \Lambda(n) \sim x$$

using the well-known fact that the Riemann zeta function does not vanish on $\Re(s) = 1$.

In a fundamental paper written in 1970, Langlands [108] outlined an approach to the Sato–Tate conjecture using the theory of automorphic forms. (It is possible that some of these ideas may have had roots in the earlier work of Sato and Tate.) To simplify matters and notation, we will give only a rough outline of this approach.

Firstly, Langlands suggested the automorphic viewpoint. Thus, the conjecture of Sato–Tate was applicable in a larger context of modular forms, or more generally, to automorphic forms on $GL(2)$. For example, one could take the celebrated Ramanujan τ-function attached to the unique newform of weight 12 and level 1, and write

$$\tau(p) = 2p^{11/2} \cos\theta_p.$$

One expects the same Sato–Tate distribution for these angles θ_p as well.

Here is a brief description of the strategy of Langlands [108]. For each natural number m, put

$$L_m(s) = \prod_v \prod_{j=0}^{m} \left(1 - \frac{\alpha_v^{m-j}\beta_v^j}{Nv^s}\right)^{-1}$$

where $\alpha_v = e^{i\theta_v}$ and $\beta_v = e^{-i\theta_v}$. Langlands indicated that the theory of automorphic forms predicts that each $L_m(s)$ should extend to an entire function. In fact, if each $L_m(s)$ extends analytically for $\Re(s) \ge 1$ and does not vanish there, then by the Tauberian theorem, we deduce for $m \ge 1$,

$$\sum_{Nv \le x} \frac{\sin(m+1)\theta_v}{\sin\theta_v} = o\big(\pi(x)\big)$$

as $x \to \infty$. Having this for each m, we will show in the next section that the Sato–Tate conjecture follows. Kumar Murty [139] showed that the non-vanishing hypothesis can be dispensed with because a very elegant argument extending the classical one of Hadamard and de la Vallée Poussin allows one to show non-vanishing from having analytic continuation to $\Re(s) \ge 1$.

In the case F is the rational number field, it is now a theorem due to Wiles and others that $L_1(s)$ is essentially the L-function attached by Hecke to a classical cusp form of weight 2. Thus, in this particular case, the Langlands conjecture is established. The non-vanishing of $L_1(s)$ on $\Re(s) = 1$ is a result due to Rankin. For $m = 2$, Rankin–Selberg theory allows one to deduce that $L_2(s)$ extends to an entire function for $\Re(s) \ge 1$. The continuation of $L_2(s)$ to the entire complex plane was established by Shimura [188] in the case $F = \mathbb{Q}$ and in the general case by

Gelbart and Jacquet [58]. In very recent work, Kim and Shahidi [96] showed that $L_3(s)$ extends to an entire function, and later, Kim showed the same for $L_4(s)$. For the cases $5 \leq m \leq 9$, Kim and Shahidi have shown that $L_m(s)$ extends to a meromorphic function for all $s \in \mathbb{C}$ which is regular for $\Re(s) \geq 1$, except in the case of $m = 9$ where $L_9(s)$ may have a pole at $s = 1$. If $L_9(s)$ were to have a pole at $s = 1$, then the Sato–Tate conjecture would be false, as we will indicate below. Thus, we can go in the reverse direction. The truth of the Sato–Tate conjecture now established implies that $L_9(s)$ does not have a pole at $s = 1$.

4 Weyl's Theorem for Compact Groups

Serre [185] gave the following reformulation of the Weyl criterion for uniform distribution in the context of a compact group. Let G be a compact group, and X its space of conjugacy classes. Let μ denote its normalized Haar measure. A sequence of elements $\{x_n\}$ with $x_n \in X$ is said to be uniformly distributed in X if for every continuous function f, we have

$$\sum_{n \leq N} f(x_n) \sim N \int_X f \, d\mu$$

as N tends to infinity.

Theorem 4.1 (Weyl's criterion for compact groups) *Let G be a compact group with Haar measure μ. A sequence $\{x_n\}$ is uniformly distributed in G if and only if*

$$\sum_{n \leq N} \chi(x_n) = o(N)$$

for every irreducible character χ of G.

The classical case in Theorem 2.1 corresponds to $G = \mathbb{R}/\mathbb{Z}$ because in this case, the irreducible characters are given by $x \mapsto e^{2\pi i m x}$.

Serre gave an interesting reformulation of this criterion in the context of L-functions. Let F be a field and for each place v of F, and let $x_v \in G$. For each irreducible representation $\rho : G \to GL_n(\mathbb{C})$, we let

$$L(s, \rho) = \prod_v \det\left(I - \rho(x_v) N v^{-s}\right)^{-1}.$$

Theorem 4.2 (Serre) *Suppose that for each irreducible non-trivial representation ρ of G, the L-function $L(s, \rho)$ extends to an analytic function for $\Re(s) \geq 1$. Then, the sequence $\{x_v\}$ is uniformly distributed in X if and only if $L(s, \rho)$ does not vanish on $\Re(s) = 1$ for all irreducible ρ.*

In the context of the Sato–Tate conjecture, one considers the group $SU(2,\mathbb{C})$ where the conjugacy classes are parameterized by

$$X_\theta = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}, \quad 0 \le \theta \le \pi.$$

The image of the Haar measure in the space of conjugacy classes of $SU(2,\mathbb{C})$ is known to be

$$\frac{2}{\pi}\sin^2\theta\, d\theta.$$

The irreducible representations of $SU(2,\mathbb{C})$ are the symmetric power representations ρ_m of the standard representation ρ_1 of $SU(2,\mathbb{C})$ into $GL(2,\mathbb{C})$. We find that $L(s,\rho_m)$ as defined above by Serre coincide with $L_m(s)$ defined in Sect. 3.

Since $\operatorname{tr}\rho_m(X_\theta) = \sin(m+1)\theta/\sin\theta$, the Sato–Tate conjecture is equivalent to the assertion

$$\sum_{N(v)\le x} \frac{\sin(m+1)\theta_v}{\sin\theta_v} = o\big(\pi_F(x)\big)$$

for each natural number m. So far, this has been established only for $m \le 8$ by the work of Kim and Shahidi [96].

5 Symmetric Power L-Series of Elliptic Curves

Let K be an algebraic number field. Let E be an elliptic curve defined over K. Let S be the (finite) set of places where E has bad reduction. For each finite place $v \notin S$ of K, we know that the number of points on E mod v is given by

$$N(v) + 1 - a_v,$$

where a_v is an integer satisfying Hasse's inequality $|a_v| \le 2N(v)^{1/2}$. As in Chap. 3, for each prime ℓ, the action of $\operatorname{Gal}(\overline{K}/K)$ on the ℓ-adic Tate module gives rise to an ℓ-adic representation

$$\rho := \rho_\ell : \operatorname{Gal}(\overline{K}/K) \to GL_2(\mathbb{Q}_\ell)$$

which is integral, that is, the characteristic polynomial of $\rho_\ell(F_v)$ (where F_v denotes the Frobenius automorphism at $v \notin S$) has integer coefficients, independent of ℓ. In fact, this characteristic polynomial is $X^2 - a_v X + N(v)$. Let us write $\alpha_v N(v)^{1/2}$, $\beta_v N(v)^{1/2}$ for the two roots of the quadratic polynomial

$$X^2 - a_v X + N(v).$$

The (partial) mth symmetric power L-function is defined as

$$L_S\big(s, \operatorname{Sym}^m \rho\big) := \prod_{v\notin S}\prod_{j=0}^{m}\big(1 - \alpha_v^j \beta_v^{m-j} N(v)^{-s}\big)^{-1}.$$

Clearly, the product converges absolutely for $\Re(s) > 1$.

In 1970, Langlands [108] suggested the existence of an automorphic representation π_m attached to $GL_{m+1}(\mathbb{A}_K)$, where $\mathbb{A}_K$ denotes the adele ring of K, such that $L_S(s, \pi_m) = L_S(s, \mathrm{Sym}^m(\rho))$ where $L_S(s, \pi_m)$ is the automorphic L-function attached to π_m with the Euler factors corresponding to the places $v \in S$ removed. In fact, it is conjectured that one can define the local factors for $v \in S$ in such a way that the completed L-function, $L(s, \mathrm{Sym}^m(\rho))$ (which we shall abbreviate as $L_m(s)$), satisfies a functional equation relating s to $1 - s$. (See, for example, [37] for details.)

If the Langlands conjecture about the existence of π_m is true, then by the theory of automorphic representations, one immediately has analytic continuation of $L_S(s, \mathrm{Sym}^m(\rho))$ to the entire complex plane. and by the result of K. Murty [139], the non-vanishing on the line $\Re(s) = 1$ follows, and the Sato–Tate conjecture follows. The non-vanishing of the L-function on the line $\Re(s) = 1$ can also be deduced from a celebrated result of Jacquet and Shalika [88], who showed that for any automorphic representation π, we have $L(s, \pi) \neq 0$ for $\Re(s) = 1$. So, what is known about the analytic continuation of $L_S(s, \mathrm{Sym}^m(\rho))$? For $m = 1$, this is the (partial) Hasse–Weil L-series attached to the elliptic curve E. It is possible to define the Euler factors at places in S as well, so that the completed L-function conjecturally admits an analytic continuation to the entire complex plane and satisfies a suitable functional equation. In the case $K = \mathbb{Q}$, the Taniyama conjecture, proved by Wiles (in the semistable case) [204] and by Breuil, Conrad, Diamond and Taylor (in the general case) [25], asserts that there is a classical cusp form f of weight 2 and level N (the conductor of E) such that the Hecke L-series $L(s, f)$ attached to f agrees with $L_1(s - 1/2)$. If π_f is the automorphic representation associated to f, then, in the context of the Langlands program, we have $L(s, \pi_f) = L_1(s)$. The series $L_1(s)$ is essentially the L-function attached to ρ, coming from the action of $\mathrm{Gal}(\overline{K}/K)$ on the Tate module.

More generally, $L_m(s)$ is essentially the L-function attached to the representation $\mathrm{Sym}^m(\rho)$, which comes from the action of $\mathrm{Gal}(\overline{K}/K)$. As alluded to above, one expects the existence of an automorphic representation π_m attached to $GL_{m+1}(\mathbb{A}_K)$ satisfying $L(s, \pi_m) = L_m(s)$. This expectation is far from being realized, though important advances have been made in this direction.

What Taylor proves is not the automorphy of $L_m(s)$ but rather its "potential automorphy." This fact, combined with other results in the analytic theory of automorphic L-functions, leads to the Sato–Tate conjecture. This result of "potential automorphy" builds on a massive collection of earlier work that can be traced back to the celebrated conjecture of Serre.

Serre [184] formulated a general conjecture about representations

$$\rho : \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to GL_2(\mathbb{F})$$

with $\mathbb{F}$ a finite field, which are odd and absolutely irreducible. More precisely, he predicted that such representations arise from classical modular forms. Serre [184] showed that his conjecture implies the Taniyama conjecture. On the other hand, Frey [55] had a remarkable insight which was completed and made more precise

by Ribet [172], who showed that the Taniyama conjecture implies Fermat's last theorem. Thus, Serre's conjecture offered a new approach to Fermat. Wiles [204] proved an important case of Serre's conjecture that was enough to deduce that of Taniyama. This work led to further developments. Most recently, Chandrashekhar Khare [93] proved the level 1 case of the Serre conjecture. In very recent work, Khare and Wintenberger [94] have settled the general case for which they were awarded the Cole Prize by the American Mathematical Society in 2011. This last development not only gives a new proof of Fermat's last theorem, but it also implies the strong Artin conjecture [9] for two-dimensional Galois representations of odd conductor.

In his recent papers, Taylor [194] made substantial progress towards this conjecture. If K is a totally real field and m is odd, he showed that there is a finite, totally real Galois extension L/K such that $(\mathrm{Sym}^m \rho)$ restricted to L is automorphic over L. One can choose an L that works simultaneously for any finite set of odd numbers. One can also choose it to be unramified at any finite set of places. Once this theorem is in hand, Taylor uses standard results from the theory of automorphic L-functions to deduce the Sato–Tate conjecture. We will give an outline of this deduction below.

The key result is the potential automorphy property of the symmetric power of a cuspidal automorphic representation, now established for a large class of such representations.

First, let us review some essential theorems from the theory of automorphic L-functions. We refer the reader to [130] for details, definitions and additional references to the literature. More precisely, we highlight pages 119 and 215 of [9] for the exact definitions of the notions of base change and automorphic induction. Here are the key theorems we will need.

First is the theorem of base change and automorphic induction, due to Arthur and Clozel [9]. This says the following. Suppose that L/K is a cyclic extension and π, Π are cuspidal representations of $GL_n(\mathbb{A}_K)$ and $GL_n(\mathbb{A}_L)$ respectively. Then, the base change of π, denoted $B(\pi)$, and the automorphic induction $I(\Pi)$ of Π exist.

The second fact we need is a celebrated theorem of Jacquet and Shalika [88], which states that for any unitary cuspidal automorphic representation π of $GL_n(\mathbb{A}_K)$, we have $L(1+it,\pi) \neq 0$.

The third fact needed is a non-vanishing theorem due to Shahidi [187]. This states that the Rankin–Selberg L-function $L(s,\pi_1 \times \pi_2)$ does not vanish on the line $\Re(s) = 1$ whenever π_1 and π_2 are unitary cuspidal automorphic representations.

A fourth fact needed is the Artin reciprocity law which states that every abelian Artin L-function of any Galois extension of K is a Hecke L-function and corresponds to a cuspidal automorphic representation of $GL_1(\mathbb{A}_K)$.

6 An Outline of the Proof of the Sato–Tate Conjecture

Here is a brief outline of the proofs in [17], [18] and [194] of the Sato–Tate conjecture. Their main theorem is: let K be a totally real field, and E/K an elliptic curve

with multiplicative reduction at some prime. For any odd number m, there is a finite, totally real Galois extension L/K such that $\mathrm{Sym}^m \rho$ becomes automorphic over L. In other words, $(\mathrm{Sym}^m \rho)|_L$ is automorphic over L. (One can also choose an L that will work simultaneously for any finite set of *odd* positive numbers.)

From this result they deduce the Sato–Tate conjecture in three steps. Here is an outline.

Step 1: For any intermediate field $K \subset F \subset L$ with L/F solvable, $\mathrm{Sym}^m \rho$ is automorphic over F. In other words, $(\mathrm{Sym}^m \rho)|_F$ is automorphic over F for every F with L/F solvable.

This is proved in his earlier paper [73].

Essentially, it applies the Arthur–Clozel theory of base change, the key idea being that the base change lift of $\mathrm{Sym}^m \rho$ to L, which exists by Taylor's main theorem, is Galois invariant and so must be the base change lift of an automorphic representation π_F for every intermediate F with L/F solvable. This is because one can find a chain

$$F = F_0 \subset F_1 \subset F_2 \subset \cdots \subset F_m = L$$

of extensions so that F_{i+1}/F_i is cyclic for $0 \leq i \leq m-1$ and apply the Arthur–Clozel theorem for automorphic induction successively, in stages, to each of the cyclic extensions $F_m/F_{m-1}, \ldots, F_1/F_0$. We refer the reader to [9] for precise details concerning automorphic induction.

Step 2: Let $G = \mathrm{Gal}(L/K)$. Now apply Brauer induction to write

$$1 = \sum_i a_i \, \mathrm{Ind}_{H_i}^G \psi_i$$

with a_i integers and ψ_i one-dimensional characters of nilpotent subgroups H_i of G. Then,

$$L\big(s, \big(\mathrm{Sym}^m \rho\big) \otimes 1\big) = \prod_i L\big(s, \big(\mathrm{Sym}^m \rho\big) \otimes \mathrm{Ind}_{H_i}^G \psi_i\big)^{a_i}.$$

By Frobenius reciprocity,

$$\big(\mathrm{Sym}^m \rho\big) \otimes \mathrm{Ind}_{H_i}^G \psi_i = \mathrm{Ind}_{H_i}^G \big(\big(\mathrm{Sym}^m \rho\big)\big|_{L^{H_i}} \otimes \psi_i\big).$$

By Step 1, $(\mathrm{Sym}^m \rho)|_{L^{H_i}}$ is automorphic over L^{H_i}. By Artin reciprocity, ψ_i is a Hecke character χ_i of L^{H_i}. Thus, $(\mathrm{Sym}^m \rho)|_{L^{H_i}} \otimes \psi_i$ is automorphic over L^{H_i}. By invariance of L-series under induction, we deduce

$$L\big(s, \mathrm{Sym}^m \rho\big) = \prod_i L\big(s, \big(\mathrm{Sym}^m \rho\big)\big|_{L^{H_i}} \otimes \chi_i\big)^{a_i},$$

and the product on the right-hand side, being a product of automorphic L-functions by Step 1, represents a meromorphic function of s. In this way, one derives the meromorphic continuation of the *odd* symmetric power L-functions attached to E.

Step 3: If we apply the Jacquet–Shalika theorem which assures us that there are no poles on $\Re(s) = 1$ for a cuspidal automorphic L-function $L(s, \pi)$, as well as the non-vanishing of $L(s, \pi)$ on $\Re(s) = 1$ for any automorphic representation π, we obtain from the above product the analytic continuation and non-vanishing on $\Re(s) = 1$ of $L(s, \mathrm{Sym}^m \rho)$ for m odd. To treat m even, one uses induction and the identity

$$\mathrm{Sym}^{m-1} \rho \oplus \mathrm{Sym}^{m+1} \rho = \mathrm{Sym}^m \rho \otimes \mathrm{Sym}^1 \rho,$$

which is essentially the trigonometric identity

$$\frac{\sin m\theta}{\sin \theta} + \frac{\sin(m+2)\theta}{\sin \theta} = \left(\frac{\sin(m+1)\theta}{\sin \theta}\right)\left(\frac{\sin 2\theta}{\sin \theta}\right)$$

(or the Clebsch–Gordon branching rule for SL_2). Thus,

$$L\big(s, \mathrm{Sym}^{m-1} \rho\big) L\big(s, \mathrm{Sym}^{m+1} \rho\big) = L\big(s, \big(\mathrm{Sym}^m \rho\big) \otimes \mathrm{Sym}^1 \rho\big).$$

Now apply Step 1 with the two odd numbers $1, m$ to get the base change to L of both $\mathrm{Sym}^m \rho$ and $\mathrm{Sym}^1 \rho$ automorphic. By the same Brauer induction trick of Step 2 applied to the right-hand side, one deduces that the right-hand side has a meromorphic continuation for all complex s. To get analytic continuation to $\Re(s) = 1$, one needs to apply Shahidi's results on the non-vanishing of Rankin–Selberg L-functions which appear on the right-hand side. Poles on the line $\Re(s) = 1$ can also be ruled out by the same theory. This completes the proof.

In many respects, this is the elliptic analogue of Brauer's theorem of the meromorphy of Artin L-series. As can be seen, the non-vanishing of the L-series on the line $\Re(s) = 1$ is essential in the proof. This was ensured by an application of theorems of Jacquet, Shalika and Shahidi. The Jacquet–Shalika theorem and the Shahidi theorem rely on the theory of Eisenstein series (à la Langlands). There are other ways of establishing non-vanishing of the L-series concerned without using the theory of Eisenstein series. Indeed, if one is willing to admit Rankin–Selberg theory and existence and analyticity of these L-functions, then classical non-vanishing techniques (of the type used by Hadamard and de la Vallée Poussin) actually work. This method is outlined in a recent paper of Sarnak [176].

7 A Chebotarev–Sato–Tate Theorem and Generalizations

In this section, we indicate the proof of Theorem 1.1. Before doing so, we discuss some natural generalizations of the Sato–Tate conjecture that one can consider. We indicate briefly what can actually be proved. Firstly, we can take two non-isogenous elliptic curves and consider the joint distribution of the angles. For instance, if E_1 and E_2 are non-isogenous elliptic curves, both without CM, defined over $\mathbb{Q}$, and $\theta_p(E_1)$ and $\theta_p(E_2)$ are the angles respectively, it is reasonable to expect that the

distribution of the pair of angles $(\theta_p(E_1), \theta_p(E_2))$ is given by the product distribution

$$\frac{4}{\pi^2} \sin^2 \theta_1 \sin^2 \theta_2 \, d\theta_1 \, d\theta_2.$$

More generally, one can consider two distinct non-CM automorphic representations and discuss their joint distribution. One would, of course, expect a similar distribution in the general case.

To prove such an assertion, we can use the formalism of Serre [185]. Using the formalism of [185], it is not difficult to show that this involves the study of certain L-series. Indeed, if both curves are non-CM and have associated Galois representations ρ_1 and ρ_2 respectively, then one needs to show that the L-series

$$L\big(s, \mathrm{Sym}^{m_1}(\rho_1) \otimes \mathrm{Sym}^{m_2}(\rho_2)\big)$$

extends to $\Re(s) \geq 1$ and does not vanish there. This looks like a difficult question to answer with the present state of knowledge. However, Harris has recently announced some progress in this direction.

If however, one of the curves has CM and corresponds to a Hecke character ψ, then one needs to study the L-series

$$L\big(s, \psi^{m_1} \otimes \mathrm{Sym}^{m_2}(\rho_2)\big)$$

and establish analytic continuation and non-vanishing in the region $\Re(s) \geq 1$. This can be done since only Hecke characters intervene, and these can be base-changed to any field by a well-known theorem of Weil [200]. It is also clear that one can take any number of CM elliptic curves and derive a similar theorem for the same reasons.

Proof of Theorem 1.1 By standard Tauberian theory, as discussed in [185], we need to show that for any irreducible representation τ of $G = \mathrm{Gal}(M/K)$, the L-function

$$L\big(s, \tau \otimes \mathrm{Sym}^m \pi\big)$$

is analytic and non-vanishing in the region $\Re(s) \geq 1$. Using Brauer induction, we write

$$\tau = \sum_i c_i \, \mathrm{Ind}_{H_i}^G \phi_i,$$

where c_i are integers, and ϕ_i is an abelian character of H_i, with H_i certain nilpotent subgroups of G. By our hypothesis, we can choose L so that L and M are disjoint and for any odd m, $\mathrm{Sym}^m \pi$ is automorphic over L. In particular, $\mathrm{Gal}(LM/L) = G$ and

$$L\big(s, \mathrm{Sym}^m \pi \otimes \tau\big) = \prod_i L\big(s, \mathrm{Sym}^m \pi \otimes \mathrm{Ind}_{H_i}^G \phi_i\big)^{c_i}.$$

Since we are viewing G as the Galois group of LM/L, we can re-write this, by Frobenius reciprocity, as

$$\prod_i L\big(s, \big(\mathrm{Ind}_{H_i}^G \big(\mathrm{Sym}^m \pi\big)\big|_{(LM)^{H_i}} \otimes \psi_i\big)\big)^{c_i},$$

where ψ_i is the Hecke character corresponding to ϕ_i via Artin reciprocity. As $(\mathrm{Sym}^m \rho)|_L$ is automorphic by our hypothesis, $(\mathrm{Sym}^m \pi)|_{LM}$ is automorphic by the theory of base change applied to the solvable extension LM/L. As in Step 1 of Taylor's theorem, we deduce that $(\mathrm{Sym}^m \pi)|_{(LM)^{H_i}}$ is automorphic over $(LM)^{H_i}$ by an application of the Arthur–Clozel theory. Since ψ_i is a Hecke character, we deduce that

$$\big(\mathrm{Sym}^m \pi\big)\big|_{(LM)^{H_i}} \otimes \psi_i$$

is automorphic over $(LM)^{H_i}$. Consequently, by the invariance of L-series under induction, we deduce that

$$L\big(s, \mathrm{Ind}_{H_i}^G \big(\big(\mathrm{Sym}^m \pi\big)\big|_{(LM)^{H_i}} \otimes \psi_i\big)\big)$$

is automorphic. Thus, it is analytic and non-vanishing for $\Re(s) \geq 1$. This proves that

$$L\big(s, \big(\mathrm{Sym}^m \pi\big) \otimes \chi\big)$$

extends to an analytic function for $\Re(s) \geq 1$ and does not vanish there. This proves the required assertion for m odd. For m even, we proceed as before, by induction to obtain the desired result. This completes the proof of Theorem 1.1. □

8 Concluding Remarks

We conclude this section with an alternate argument in the case that M/K is a nilpotent Galois extension which is simpler. In future variations, this alternate argument may be useful.

By standard Tauberian theory, as discussed in [185], we need to show that for any irreducible representation τ of $\mathrm{Gal}(\overline{K}/K)$, with nilpotent image, the L-function

$$L\big(s, \tau \otimes \mathrm{Sym}^m \pi\big)$$

is analytic and non-vanishing in the region $\Re(s) \geq 1$. Since any irreducible representation of a finite nilpotent group is induced from an abelian character χ of a subgroup H, so

$$\tau \otimes \mathrm{Sym}^m(\pi) = \mathrm{Ind}_H^G\big(\chi \otimes \mathrm{Sym}^m(\pi)|_{L^H}\big).$$

By Arthur–Clozel, $\chi \otimes \mathrm{Sym}^m(\pi)|_{L^H}$ is automorphic. We now complete the proof as before. It is clear from the preceding arguments that if one had the automorphic

induction of Hecke characters, the proof would go through for any Galois setting and not just in the nilpotent or solvable setting. Future advances in the Langlands program should translate into general theorems of Chebotarev–Sato–Tate type.

Just as in the classical case of the prime number theorem, one can ask about what error terms we may expect. It may be too early here to make statements that arise from the recent progress on potential automorphy. One would need more precise information about these fields $K_{S,T}$ before such error terms can be written down.

There is an interesting conjecture due to Akiyama and Tanigawa [5] predicting that for any $\epsilon > 0$,

$$\frac{\#\{p \leq x : \theta_p \in [\alpha, \beta]\}}{\#\{p \leq x\}} = \frac{2}{\pi} \int_{\alpha}^{\beta} \sin^2 \theta \, d\theta + O\left(x^{-1/2+\epsilon}\right). \tag{55}$$

But this was already proved by Kumar Murty [140] in 1985 assuming that the $L_m(s)$ satisfy the analogue of the Riemann hypothesis. In fact, this result makes explicit the dependence of the error term on the length of the interval $[\alpha, \beta]$. One can also go in the reverse direction. Assertion (55) is already a sufficiently strong conjecture, and one can use standard methods to deduce that it implies that every $L_m(s)$ satisfies the analogue of the Riemann hypothesis for $m \geq 1$.

ERRATUM

Erratum to: The Ramanujan τ-Function

Erratum to: M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*, pp. 11–23,
DOI 10.1007/978-81-322-0770-2_2,

Page 17: The third displayed equation should be a congruence (mod 2). It is clearer if we note $(1-q^n)^{24} \equiv (1-q^{8n})^3 \pmod 2$ and observe that

$$\prod_{n=1}^{\infty}\left(1-q^{8n}\right)^3 = \sum_{m=0}^{\infty}(-1)^m(2m+1)q^{4m^2+4m} \equiv q^{4m^2+4m} \pmod 2.$$

The online version of the original chapter can be found at doi: 10.1007/978-81-322-0770-2_2.

References

1. S.D. Adhikari, A. Mukhopadhyay, Partition function congruences: some flowers and seeds from "Ramanujan's garden". Expo. Math. **19**, 193–201 (2001)
2. S. Ahlgren, Distribution of parity of the partition function in arithmetic progressions. Indag. Math. (N.S.) **10**(2), 173–181 (1999)
3. S. Ahlgren, M. Boylan, Arithmetic properties of the partition function. Invent. Math. **153**(3), 487–502 (2003)
4. S. Ahlgren, M. Boylan, Odd coefficients of weakly holomorphic modular forms. Math. Res. Lett. **15**(3), 409–418 (2008)
5. S. Akiyama, Y. Tanigawa, Calculation of values of L-functions associated to elliptic curves. Math. Comp. **68**(227), 1201–1231 (1999)
6. L. Alaoglu, P. Erdös, On highly composite and similar numbers. Trans. Am. Math. Soc. **58**, 448–469 (1944)
7. G.E. Andrews, An introduction to Ramanujan's "lost" notebook. Am. Math. Mon. **86**, 89–108 (1979)
8. G.E. Andrews, F. Garvan, Dyson's crank of a partition. Bull. Am. Math. Soc. **18**, 167–171 (1988)
9. J. Arthur, L. Clozel, *Simple Algebras, Base Change and the Advanced Theory of the Trace Formula*. Annals of Mathematics Studies, vol. 120 (Princeton University Press, Princeton, 1989)
10. E. Artin, Quadratische Körper im Gebiete der Höheren kongruenzen, I, II. Math. Z. **19**, 153–246 (1924)
11. A.O.L. Atkin, Proof of a conjecture of Ramanujan. Glasg. Math. J. **8**, 14–32 (1967)
12. A.O.L. Atkin, J. Lehner, Hecke operators on $\Gamma_0(m)$. Math. Ann. **185**, 134–160 (1970)
13. A.O.L. Atkin, H.P.F. Swinnerton-Dyer, Some properties of partitions. Proc. Lond. Math. Soc. **4**, 84–106 (1954)
14. R. Balasubramanian, J.-M. Deshouillers, F. Dress, Problème de Waring pour les bicarrés, I, Schéma de la solution. C. R. Acad. Sci. Paris Sér. I Math. **303**(4), 85–88 (1986)
15. R. Balasubramanian, J.-M. Deshouillers, F. Dress, Problème de Waring pour les bicarrés, II, Résultats axillaires pour le théorème asymptotique. C. R. Acad. Sci. Paris Sér. I Math. **303**(5), 161–163 (1986)
16. R. Balasubramanian, K. Ramachandra, The place of an identity of Ramanujan in prime number theory. Proc. Indian Acad. Sci. **83A**(4), 156–165 (1976)
17. T. Barnet-Lamb, D. Geraghty, T. Gee, The Sato–Tate conjecture for Hilbert modular forms. J. Am. Math. Soc. **24**(2), 411–469 (2011)
18. T. Barnet-Lamb, D. Geraghty, M. Harris, R. Taylor, A family of Calabi–Yau varieties and potential automorphy, II. Publ. Res. Inst. Math. **47**, 29–98 (2011)
19. B.C. Berndt, *Ramanujan's Notebooks, Part II* (Springer, Berlin, 1989)

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2, © Springer India 2013

20. B.C. Berndt, *Ramanujan's Notebooks, Part V* (Springer, Berlin, 1998)
21. B.C. Berndt, Ramanujan's congruences for the partition function modulo 5, 7 and 11. Int. J. Number Theory **3**(3), 349–354 (2007)
22. B.C. Berndt, Ramanujan reaches his hand from his grave to snatch your theorems from you. Asia Pac. Math. Newsl. **1**(2), 8–13 (2011)
23. B.C. Berndt, The chief-accountant and mathematical friend of Ramanujan, S. Narayana Aiyar. Am. Math. Mon. **118**(9), 767–776 (2011)
24. R. Blecksmith, J. Brillhart, I. Gerst, Parity results for certain partition functions and identities similar to theta function identities. Math. Comp. **48**(177), 29–38 (1987)
25. C. Breuil, B. Conrad, F. Diamond, R. Taylor, On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercises. J. Am. Math. Soc. **14**(4), 843–939 (2001)
26. J. Bruinier, K. Ono, Heegner divisors, L-functions and harmonic weak Maass forms. Ann. of Math. **172**, 2135–2181 (2010)
27. J. Bruinier, K. Ono, R. Rhoades, Differential operators for harmonic weak Maass forms and the vanishing of Hecke eigenvalues. Math. Ann. **342**(3), 673–693 (2008)
28. J.H. Bruinier, K. Ono, An algebraic formula for the partition function (in press)
29. J.H. Bruinier, K. Ono, Algebraic formulas for the coefficients of half-integral weight harmonic weak Maass forms (in press)
30. H. Carayol, La conjecture de Sato–Tate (d'après Clozel, Harris, Shepherd-Barron, Taylor), Sem. Bourbaki **59**(977) (2006–2007)
31. S. Chandrasekhar, On Ramanujan, in *Ramanujan Revisited*, ed. by G.E. Andrews, et al. (Academic Press, San Diego, 1988)
32. K. Chandrasekharan, R. Narasimhan, Functional equations with multiple gamma factors and the average order of arithmetical functions. Ann. of Math. **76**(2), 93–136 (1962)
33. S. Chowla, Congruence properties of partitions. J. Lond. Math. Soc. **9**, 247 (1934)
34. S. Chowla, A. Selberg, On Epstein's zeta function. J. Reine Angew. Math. **227**, 86–110 (1967)
35. L. Clozel, M. Harris, R. Taylor, Automorphy of some ℓ-adic lifts of automorphic mod ℓ Galois representations. Publ. Math. IHES **108**(1), 1–182 (2008)
36. J. Cogdell, H. Kim, M.R. Murty, *Lectures on Automorphic L-functions* (Am. Math. Soc., Providence, 2004)
37. J. Cogdell, P. Michel, On the complex moments of symmetric power L-functions at $s = 1$. Int. Math. Res. Not. **31**, 1561–1617 (2004)
38. P. Deligne, Formes modulaires et représentations ℓ-adiques, in *Séminaire Bourbaki, Exposé 355*. Lecture Notes in Mathematics, vol. 179 (Springer, Berlin, 1971)
39. P. Deligne, La conjecture de Weil, I. Publ. Math. IHES **43**, 273–307 (1974)
40. P. Deligne, La conjecture de Weil, II. Publ. Math. IHES **52**, 138–252 (1980)
41. M. Deuring, Die Typen der Multiplikatorenringe elliptischer Funktionenkörper. Abh. Math. Zem. Hansischen Univ. **14**, 197–272 (1941)
42. M. Dewar, M.R. Murty, A derivation of the Hardy–Ramanujan formula from an arithmetic formula. Proc. Amer. Math. Soc. (in press)
43. W. Duke, Continued fractions and modular functions. Bull. Am. Math. Soc. **42**(2), 137–162 (2005)
44. D. Duverney, K. Nishioka, K. Nishioka, I. Shiokawa, Transcendence of Jacobi's theta series and related results, in *Number Theory*, Eger, 1996 (de Gruyter, Berlin, 1998), pp. 157–168
45. F. Dyson, A walk through Ramanujan's garden, in *Ramanujan Revisited* (Academic Press, Boston, 1988), pp. 7–28
46. P.D.T.A. Elliott, *Probabilistic Number Theory, I, II* (Springer, New York, 1980)
47. W.J. Ellison, Waring's problem. Am. Math. Mon. **78**, 10–36 (1971)
48. P. Erdös, On an elementary proof of some asymptotic formulas in the theory of partitions. Ann. of Math. **43**(2), 437–450 (1942)
49. P. Erdös, M. Kac, The Gaussian law of errors in the theory of additive number theoretic functions. Am. J. Math. **62**(1), 738–742 (1940)

50. P. Erdös, C. Pomerance, On the normal number of prime factors of $\phi(n)$. Rocky Mt. J. Math. **15**, 343–352 (1985)
51. T. Estermann, On the representation of a number as the sum of two products. Proc. Lond. Math. Soc. **31**(2), 123–133 (1930)
52. L. Euler, *Opera Postuma*, vol. 1, (1862), pp. 203–204
53. A. Folsom, Z.A. Kent, K. Ono, ℓ-adic properties of the partition function. Adv. Math. **229**, 1586–1609 (2012)
54. G. Frei, P. Roquette, *Emil Artin and Helmut Hasse, Die Korrespondenz 1923–1934* (Universitätsverlag, Göttingen, 2008)
55. G. Frey, Links between solutions of $A - B = C$ and elliptic curves. Lect. Notes Math. **1380**, 31–62 (1989)
56. F. Garvan, New combinatorial interpretations of Ramanujan's partition congruences mod 5, 7, and 11. Trans. Am. Math. Soc. **305**, 47–77 (1988)
57. F. Garvan, D. Kim, D. Stanton, Cranks and t-cores. Invent. Math. **101**, 1–17 (1990)
58. S. Gelbart, H. Jacquet, A relation between automorphic forms on $GL(2)$ and $GL(3)$. Proc. Natl. Acad. Sci. USA **73**, 3348–3350 (1976)
59. E. Ghate, On monomial relations between Eisenstein series. J. Ramanujan Math. Soc. **15**(2), 71–79 (2000)
60. C.C. Grosjean, Disproving a conjecture about an asymptotic formula of Ramanujan involving the divisor sums. Simon Stevin **53**, 39–70 (1979)
61. B. Gross, On an identity of Chowla and Selberg. J. Number Theory **11**(3), 344–348 (1979). S. Chowla Anniversary issue
62. E. Grosswald, Die Werte der Riemannschen Zetafunktion an ungeranden Argumentstellen. Nachr. Akad. Wiss., Gottinger Math.-Phys. Kl. **2**, 9–13 (1970)
63. A. Grothendieck, The cohomology theory of abstract algebraic varieties. Proc. Int. Congr. Math. Edinburgh, 103–118 (1958)
64. S. Gun, M.R. Murty, P. Rath, Algebraic independence of values of modular forms. Int. J. Number Theory **7**, 1065–1074 (2011)
65. S. Gun, M.R. Murty, P. Rath, Transcendental values of certain Eichler integrals. Bull. Lond. Math. Soc. **43**(5), 939–952 (2011)
66. H. Halberstam, An asymptotic formula in the theory of numbers. Trans. Am. Math. Soc. **84**, 338–351 (1957)
67. G.H. Hardy, *Ramanujan: Twelve Lectures on Subjects Suggested by his Life and Work*, 3rd edn. (Chelsea, New York, 1978)
68. G.H. Hardy, J.E. Littlewood, Contributions to the theory of the Riemann zeta function and the distribution of primes. Acta Math. **41**, 119–196 (1918)
69. G.H. Hardy, S. Ramanujan, The normal number of prime factors of a number n. Q. J. Math. **48**, 76–92 (1917)
70. G.H. Hardy, S. Ramanujan, Asymptotic formulae for the distribution of integers of various types. Proc. Lond. Math. Soc. **16**(2), 112–132 (1917)
71. G.H. Hardy, S. Ramanujan, Asymptotic formulae in combinatory analysis. Proc. Lond. Math. Soc. **17**(2), 75–115 (1918)
72. G.H. Hardy, E.W. Wright, *Introduction to the Theory of Numbers*, 5th edn. (Oxford University Press, London, 1979)
73. M. Harris, N. Shepherd-Barron, R. Taylor, A family of Calabi–Yau varieties and potential automorphy. Preprint available at www.math.harvard.edu/~rtaylor
74. H. Hasse, Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F.K. Shmidtschen Kongruenzzetafunktionen in gewissen zykischen Fällen, Vorläufige Mitteilung. Nachr. Ges. Wiss. Göttingen I. Math.-Phys. Kl. Fachgr. I Math. **42**, 253–262 (1933)
75. E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzalhen, I. Math. Z. **1**, 357–376 (1918)
76. E. Hecke, Eine neue Art von Zetafunktionen und ihre Beziehungen zur Verteilung der Primzahlen, II. Math. Z. **6**, 11–51 (1920)

77. E. Hecke, Die primzahlen in der theorie der elliptischen modulfunktionen, in *Mathematische Werke*, Paper 32 (Vandenhoeck & Ruprecht, Gortingen, 1983), p. 577
78. D.R. Hickerson, A proof of the mock theta conjectures. Invent. Math. **94**, 639–660 (1988)
79. D.R. Hickerson, On the seventh order mock theta functions. Invent. Math. **94**, 661–677 (1988)
80. D. Hilbert, Beweis für die Darstellbarkeit der ganzen Zahlen durn eine feste Anzahl n-ter Potenzen (Waringsches Problem). Math. Ann. **67**, 281–300 (1909)
81. M.D. Hirschhorn, M.V. Subbarao, On the parity of $p(n)$. Acta Arith. **50**(4), 351–355 (1988)
82. G. Hoheisel, Primzahlproblem in der Analysis. Berlin Math. Ges. Sitzungsber, 550–558 (1930)
83. L.K. Hua, *Introduction to Number Theory* (Springer, Berlin, 1982)
84. A.E. Ingham, Some asymptotic formulae in the theory of numbers. Journal of the London Mathematical Society, 202–208 (1927)
85. A.E. Ingham, A note on Riemann's ζ-function and Dirichlet's L-functions. J. Lond. Math. Soc. **5**, 107–112 (1930)
86. H. Jacquet, R.P. Langlands, *Automorphic Forms on GL*(2). Lecture Notes in Mathematics, vol. 114 (Springer, Berlin, 1970)
87. H. Jacquet, I.I. Piatetski-Shapiro, J. Shalika, Rankin–Selberg convolutions. Am. J. Math. **105**, 367–464 (1983)
88. H. Jacquet, J. Shalika, A non-vanishing theorem for zeta functions of GL_n. Invent. Math. **38**(1), 1–16 (1976/77)
89. H. Jacquet, J. Shalika, On Euler products and the classification of automorphic representations I. Am. J. Math. **103**, 499–558 (1981)
90. M. Kaneko, D. Zagier, A generalized Jacobi theta function and quasimodular forms, in *The Moduli Space of Curves*. Progress in Mathematics, vol. 129 (Birkhäuser, Boston, 1995), pp. 165–172
91. N. Katz, An overview of Deligne's proof of the Riemann hypothesis for varieties over finite fields. Proc. Symp. Pure Math. **28**, 275–305 (1976)
92. A. Kempner, Bemerkungen zum Waringschen Problem. Math. Ann. **72**(3), 387–399 (1912)
93. C. Khare, Serre's modularity conjecture: the level one case. Duke Math. J. **134**(3), 557–589 (2006)
94. C. Khare, J.-P. Wintenberger, Serre's modularity conjecture, I and II. Invent. Math. **178**(3), 485–504 (2009) (and 505–586)
95. H. Kim, P. Sarnak, Refined estimates towards the Ramanujan and Selberg conjectures in the appendix to functoriality for the exterior square of GL_4 and the symmetric fourth of GL_2. J. Am. Math. Soc. **16**(1), 139–183 (2003)
96. H. Kim, F. Shahidi, Functorial products for $GL_2 \times GL_3$ and functorial symmetric cube for GL_2. C. R. Acad. Sci. Paris Sér. I Math. **331**(8), 599–604 (2000)
97. H. Kim, F. Shahidi, Functorial products for $GL_2 \times GL_3$ and the symmetric cube for GL_2. Ann. of Math. (2) **155**(3), 837–893 (2002).
98. H.D. Kloosterman, Asymptotische Formeln für die Fourierkoeffizienten ganzer Modulformen. Abh. Math. Semin. Univ. Hamb. **5**, 338–352 (1927)
99. S. Knapowski, P. Turan, Comparative prime number theory, I–VIII. Acta Math. Hung. **13**, 299–364 (1962)
100. O. Kölberg, Note on the parity of the partition function. Math. Scand. **7**, 377–378 (1959)
101. C. Krattenthaler, T. Rivoal, W. Zudilin, Séries hypergéométriques basiques, q analogues des valeurs de la fonction zeta et séries d'Eisenstein. J. Inst. Math. Juissieu **5**(1), 53–79 (2006)
102. N.V. Kuznetsov, The Petersson conjecture for cusp forms of weight zero and the Linnik conjecture: sums of Kloosterman sums. Mat. Sb. (N.S.) **111(153)**(3), 334–383 (1980), 479 (in Russian)
103. J. Lagarias, An elementary problem equivalent to the Riemann hypothesis. Am. Math. Mon. **109**, 534–543 (2002)
104. D.B. Lahiri, On Ramanujan's function $\tau(n)$ and the divisor function $\sigma_k(n)$, I. Bull. Calcutta Math. Soc. **38**, 193–206 (1946). II, **39**, 33–52 (1947)

105. E. Landau, Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie. Math. Zeit. **1**, 1–24 (1918) (and 213–219)
106. S. Lang, *Introduction to Modular Forms* (Springer, Berlin, 1976)
107. S. Lang, *Elliptic Functions*, 2nd edn. (Springer, New York, 1987)
108. R. Langlands, Problems in the theory of automorphic forms, in *Lectures in Modern Analysis and Applications*, ed. by R. Dudley, Springer Lecture Notes in Mathematics, vol. 170 (1970), pp. 18–86
109. D. Lanphier, Maass operators and van der Pol-type identities for Ramanujan's tau function. Acta Arith. **113**, 157–167 (2004)
110. D.H. Lehmer, Ramanujan's function $\tau(n)$. Duke Math. J. **10**, 483–492 (1943)
111. D.H. Lehmer, The vanishing of Ramanujan's function $\tau(n)$. Duke Math. J. **14**, 429–433 (1947)
112. J.E. Littlewood, Distribution des nombres premiers. C. R. Acad. Sci. Paris **158**, 1869–1872 (1914)
113. Y.R. Liu, M.R. Murty, The Turán sieve method and some of its applications. J. Ramanujan Math. Soc. **14**(1), 21–35 (1999)
114. Y.R. Liu, M.R. Murty, Sieve methods in combinatorics. J. Comb. Theory, Ser. A **111**(1), 1–23 (2005)
115. Y.R. Liu, M.R. Murty, A weighted Turán sieve method. J. Number Theory **116**(1), 1–20 (2006)
116. N. Lygeros, O. Rozier, A new solution to the equation $\tau(p) \equiv 0 \pmod{p}$. J. Integer Seq. **13**(7) (2010). Article 10.7.4, 11 pp
117. H. Maass, Über eine neue Art von nichtanalytischen automorphen Funktionen und die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen. Math. Ann. **121**, 141–182 (1944)
118. H. Maass, *Siegel Modular Forms and Dirichlet Series*. Lecture Notes in Mathematics, vol. 216 (Springer, Berlin, 1971)
119. K. Mahler, On the fractional parts of the powers of a rational number II. Mathematika **4**, 122–124 (1957)
120. H.B. Mann, A proof of the fundamental theorem on the density of sums of sets of positive integers. Ann. of Math. **43**(3), 523–527 (1942)
121. F. Martin, E. Royer, Formes modulaires et périodes, in *Formes modulaires et transcendance*, Sémin. Congr., vol. 12 (Soc. Math. France, Paris, 2005), pp. 1–17
122. D.W. Masser, *Elliptic Functions and Transcendence*. Lecture Notes in Mathematics, vol. 437 (Springer, Berlin, 1975)
123. B. Mazur, Controlling our errors. Nature **443**, 38–40 (2006)
124. B. Mazur, Finding meaning in error terms. Bull. Am. Math. Soc. **45**, 185–228 (2008)
125. C. Moeglin, J.-L. Waldspurger, Le spectre résiduel de $GL(n)$. Ann. Sci. Éc. Norm. Super. (4) **22**, 605–674 (1989)
126. L.J. Mordell, On Ramanujan's empirical expansions of modular functions. Proc. Camb. Philos. Soc. **19**, 117–124 (1920)
127. Y. Motohashi, The binary additive divisor problem. Ann. Sci. Éc. Norm. Super. **27**, 529–572 (1994)
128. M.R. Murty, Oscillations of Fourier coefficients of modular forms. Math. Ann. **262**, 431–446 (1983)
129. M.R. Murty, On the estimation of eigenvalues of Hecke operators. Rocky Mt. J. Math. **15**(2), 521–533 (1985)
130. M.R. Murty, Recent developments in the Langlands program. C. R. Math. Rep. Sci. Canada **24**(2), 33–54 (2002)
131. M.R. Murty, *Problems in Analytic Number Theory*, 2nd edn. (Springer, New York, 2005)
132. M.R. Murty, V.K. Murty, Prime divisors of Fourier coefficients of modular forms. Duke Math. J. **51**, 57–76 (1984)
133. M.R. Murty, V.K. Murty, An analogue of the Erdös–Kac theorem for Fourier coefficients of modular forms. Indian J. Pure Appl. Math. **15**, 1090–1101 (1984)

134. M.R. Murty, V.K. Murty, T.N. Shorey, Odd values of the Ramanujan τ-function. Bull. Soc. Math. Fr. **115**(3), 391–395 (1987)
135. M.R. Murty, P. Rath, Introduction to transcendental number theory (in press)
136. M.R. Murty, F. Saidak, Non-abelian generalizations of the Erdös–Kac theorem. Can. J. Math. **56**(2), 356–372 (2004)
137. M.R. Murty, C. Smyth, R. Wang, Zeros of Ramanujan polynomials. J. Ramanujan Math. Soc. **26**, 107–125 (2011)
138. M.R. Murty, C. Weatherby, Special values of the gamma function at CM points. Int. J. Number Theory (in press)
139. V.K. Murty, On the Sato–Tate conjecture, in *Number Theory Related to Fermat's Last Theorem*. Cambridge, MA, 1981. Progress in Mathematics, vol. 26 (Birkhäuser, Boston, 1982), pp. 195–205
140. V.K. Murty, Explicit formulae and the Lang–Trotter conjecture. Rocky Mt. J. Math. **15**(2), 535–551 (1985)
141. V.K. Murty, Ramanujan and Harish-Chandra. Math. Intell. **15**(2), 33–39 (1993)
142. Yu.V. Nesterenko, Algebraic independence for values of Ramanujan functions, in *Introduction to Algebraic Independence Theory*. Lecture Notes in Math., vol. 1752 (Springer, Berlin, 2001), pp. 27–46
143. D.J. Newman, Evaluation of the constant in the formula for the number of partitions of n. Am. J. Math. **73**(3), 599–601 (1951)
144. M. Newman, Periodicity modulo m and divisibility properties of the partition function. Trans. Am. Math. Soc. **97**(2), 225–236 (1960)
145. J.L. Nicolas, Répartition des nombres hautement composés de Ramanujan. Can. J. Math. **23**, 116–130 (1971)
146. J.L. Nicolas, Grandes valeurs des fonctions arithmétiques. Sém. Delange-Pisot-Poitou, Paris, no. G20, 5 p. (1974/75)
147. J.L. Nicolas, On highly composite numbers, in *Ramanujan Revisited*, ed. by G. Andrews, et al. (Academic Press, San Diego, 1988), pp. 215–244
148. J.L. Nicolas, I.Z. Rusza, A. Sarkozy, Partition function (with an appendix by J.-P. Serre). J. Number Theory **73**, 292–317 (1998)
149. D. Niebur, A formula for Ramanujan's τ-function. Ill. J. Math. **19**, 448–449 (1975)
150. K. Ono, Parity of the partition function in arithmetic progressions. J. Reine Angew. Math. **472**, 1–15 (1996)
151. K. Ono, *The Web of Modularity: Arithmetic of the Coefficients of Modular Forms and q-Series*. CBMS Regional Conference Series in Mathematics, vol. 102 (Am. Math. Soc., Providence, 2004)
152. K. Ono, Unearthing the visions of a master, in *Harmonic Maass forms and Number Theory*. Current Developments in Mathematics, (2008), pp. 347–454
153. T.R. Parkin, D. Shanks, On the distribution of parity in the partition function. Math. Comp. **21**, 466–480 (1967)
154. H. Petersson, Theorie der automorphen formen beliebiger reeler Dimension und ihre Darstellung durch eine neue art Poincarészcher Reihen. Math. Ann. **103**, 346–369 (1930)
155. S.S. Pillai, On Waring's problem $g(6) = 73$. Proc. Indian Acad. Sci. **12A**, 30–40 (1940)
156. S.S. Pillai, *Collected Papers*, vol. 1, ed. by R. Balasubramanian, R. Thangadurai (Ramanujan Mathematical Society, India, 2010), p. xiii
157. H. Rademacher, On the expansion of the partition function in a series. Ann. of Math. **44**, 416–422 (1943)
158. S. Ramanujan, On the product $\prod_{n=0}^{\infty}[1 + (\frac{x}{a+nd})^3]$. J. Indian Math. Soc. **7**, 209–211 (1915)
159. S. Ramanujan, Some definite integrals. Messenger of Math. **44**, 10–18 (1915)
160. S. Ramanujan, Highly composite numbers. Proc. Lond. Math. Soc. **14**(2), 347–400 (1915)
161. S. Ramanujan, On certain arithmetical functions. Trans. Cambridge Philos. Soc. **22**, 159–184 (1916)

162. S. Ramanujan, On certain arithmetical functions. Trans. Camb. Phil. Soc. **22**, 159–184 (1916). See also *Collected Papers*, Number 18 (AMS–Chelsea, Providence, 1962), pp. 136–192
163. S. Ramanujan, Proof of certain identities in combinatory analysis. Proc. Camb. Philos. Soc. **19**, 214–216 (1919)
164. S. Ramanujan, Some properties of $p(n)$, the number of partitions of n. Proc. Camb. Philos. Soc. **19**, 207–210 (1919)
165. S. Ramanujan, Congruence properties of partitions. Math. Z. **9**, 147–153 (1921)
166. S. Ramanujan, Highly composite numbers, annotated and with a foreword by Jean-Louis Nicolas and Guy Robin. Ramanujan J. **1**(2), 119–153 (1997)
167. B. Ramakrishnan, B. Sahu, Rankin–Cohen brackets and van der Pol-type identities for the Ramanujan τ-function. Preprint (2007)
168. O. Ramaré, On Snirel'man's constant. Ann. Sc. Norm. Super. Pisa Cl. Sci. (4), **22**(4), 645–706 (1995)
169. R.A. Rankin, Contributions to the theory of Ramanujan's function $\tau(n)$ and similar arithmetical functions, II, the order of the Fourier coefficients of integral modular forms. Proc. Camb. Philos. Soc. **35**, 357–372 (1939)
170. R.A. Rankin, The construction of automorphic forms from the derivatives of a given form. J. Indian Math. Soc. **20**, 103–116 (1956)
171. R.A. Rankin, *Modular Forms and Functions* (Cambridge University Press, Cambridge, 1977)
172. K. Ribet, From the Taniyama–Shimura conjecture to Fermat's last theorem. Ann. Fac. Sci. Toulouse Ser. 5 **11**(1), 116–139 (1990)
173. T. Rivoal, La fonction zeta de Riemann prend une infinité de valeurs irrationelles aux entiers impairs. C. R. Acad. Sci. Paris Sér. I Math. **331**(4), 267–270 (2000)
174. G. Robin, Estimation de la fonction de Tchebycheff θ sur le k-ieme nombre premier et grandes valeurs de la fonction $\omega(n)$ nombre de diviseurs premiers de n. Acta Arith. **42**, 367–389 (1983)
175. L.J. Rogers, Second memoir on the expansion of certain infinite products. Proc. Lond. Math. Soc. **25**(1), 318–343 (1894)
176. P. Sarnak, Non-vanishing of L-functions on $\Re(s) = 1$, in *Contributions to Automorphic Forms, Geometry, and Number Theory* (Johns Hopkins University Press, Baltimore, 2004), pp. 719–732
177. I. Schur, Ein Beitrag zur additiven Zahlentheorie und zur Theorie der Kettenbruche. Berliner Sitzungsber. **23**, 301–321 (1917)
178. A. Selberg, Über einge arithmetische identitäten. Avh. Norske Vidensk. Akad. Oslo **1**(8), (1936), 23s
179. A. Selberg, Über die Fourierkoeffizienten elliptischer Modulformen negativer Dimension, Neuvième Congrès des Mathematiciens Scandinaves, Helsingfors, 1938 (1939), pp. 320–322
180. A. Selberg, On the estimation of Fourier coefficients of modular forms, in *Proc. Symp. Pure Math.*, vol. 8 (Amer. Math. Soc., Providence, 1965), pp. 1–15
181. A. Selberg, Reflections around the Ramanujan centenary, in *Collected Papers*, vol. 1 (Springer, Berlin, 1989)
182. J.-P. Serre, Zeta and L-functions, in *Arithmetic Algebraic Geometry*. Proc. Conf. Purdue University (Harper & Row, New York, 1963), pp. 82–92
183. J.-P. Serre, *A Course in Arithmetic*. Graduate Texts in Mathematics, vol. 7 (Springer, New York, 1973)
184. J.-P. Serre, Sur les répresentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Duke Math. J. **54**, 179–230 (1987)
185. J.-P. Serre, *Abelian ℓ-Adic Representations and Elliptic Curves*. Research Notes in Mathematics, vol. 7 (AK Peters, Wellesley, 1998)
186. I.R. Shafarevich, On certain tendencies in the development of mathematics. Math. Intell. **3**(4), 182–184 (1980/81)

187. F. Shahidi, On certain L-functions. Am. J. Math. **103**, 297–355 (1981)
188. G. Shimura, On the holomorphy of certain Dirichlet series. Proc. Lond. Math. Soc. (3) **31**(1), 79–98 (1975)
189. G. Shimura, The special values of the zeta functions associated with cusp forms. Commun. Pure Appl. Math. **29**, 783–804 (1976)
190. G. Shimura, Differential operators, holomorphic projections, and singular forms. Duke Math. J. **76**, 141–173 (1994)
191. T. Shioda, On elliptic modular surfaces. J. Math. Soc. Jpn. **24**, 20–59 (1972)
192. M.V. Subbarao, Some remarks on the partition function. Am. Math. Mon. **73**, 851–854 (1966)
193. J. Tate, Algebraic cycles and poles of zeta functions, in *Arithmetic Algebraic Geometry*, ed. by F.G. Schilling (Harper & Row, New York, 1965), pp. 93–110
194. R. Taylor, Automorphy of some ℓ-adic lifts of automorphic mod ℓ representations, II. Publ. Math. IHES **108**(1), 183–239 (2008)
195. P.L. Tchebycheff, Lettre de M. le professeur Tchebycheff à M. Fuss sur un nouveau théorème relatif aux nombres premiers dans la forme $4n+1$ et $4n+3$. Bull. Cl. Phys.-Math. Acad. St. Petersburg **11**, 208 (1853)
196. P. Turan, On a theorem of Hardy and Ramanujan. J. Lond. Math. Soc. **9**, 274–276 (1934)
197. R.C. Vaughan, *The Hardy–Littlewood Method*, 2nd. edn. Cambridge Tracts in Mathematics (Cambridge University Press, Cambridge, 1997)
198. A. Walfisz, *Weylsche Exponential Summen in der Neueren Zahlentheorie* (VEB Deutscher Verlag der Wissenschaften, Berlin, 1963)
199. G.N. Watson, Ramanujan's vermutung über zerfällungsanzahlen. J. Reine Angew. Math. **179**, 97–128 (1938)
200. A. Weil, *Dirichlet Series and Automorphic Forms*. Springer Lecture Notes, vol. 189, (Springer, Berlin, 1971)
201. A. Weil, Two lectures on number theory, past and present. Enseign. Math. **20**, 87–110 (1974)
202. A. Wieferich, Beweis des Satzes, das sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen lässt. Math. Ann. **66**(1), 95–101 (1909)
203. S. Wigert, Sur l'ordre de grandeur du nombre de diviseurs d'un entire. Ark. Mat. **3**(18), 1–9 (1907)
204. A. Wiles, Modular elliptic curves and Fermat's last theorem. Ann. of Math. **141**(3), 443–551 (1995)
205. H. Yoshida, An analogue of the Sato–Tate conjecture. Invent. Math. **19**, 261–277 (1973)
206. D. Zagier, Ramanujan's mock theta functions and their applications, (d'après Zwegers and Bringmann-Ono). Sem. Bourbaki **60**(986) (2007)
207. D. Zagier, Elliptic Modular Forms and Their Applications, in *The 1–2–3 of Modular Forms*. Universitext (Springer, Berlin, 2008), pp. 1–103

Index

M.R. Murty, V.K. Murty, *The Mathematical Legacy of Srinivasa Ramanujan*,
DOI 10.1007/978-81-322-0770-2, © Springer India 2013